2018

Phenotypic and genomic selection for multi-trait improvement in soybean line and variety development

Andreomar Kurek
Iowa State University

Phenotypic and genomic selection for multi-trait improvement in soybean line and variety development

by

Andreomar José Kurek

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Plant Breeding

Program of Study Committee:
William Beavis, Major Professor
Alicia Carriquiry
Dorian Garrick
Lizhi Wang
Silvia Cianzio
Thomas Lubberstedt

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2018

TABLE OF CONTENTS

Page

LIST OF FIGURES x
LIST OF TABLES xii
NOMENCLATURE xvi
ABSTRACT xvii
INTRODUCTION 1
METHODS 16
Variety development 16
Genotypic data from potential founders 17
Simulated genomes of progeny in filial generations evaluated at each stage of development 18
Simulated phenotypic values 19
Simulation experiments 21
Evaluation metrics 26
RESULTS 30
Genotypic data 30
Simulated phenotypes of the founders and their progeny 31
Simulation experiment 1: phenotypic selection 32
Simulation experiment 2: MAS for SCN and PRR traits + phenotypic selection for yield 36
Simulation experiment 3: Selections of crosses based on performance of 112 potential founders + MAS for SCN and PRR traits in early generations + phenotypic selection for yield through all stages of development. 40
Simulation experiment 4: Selections of crosses based on performance of lines in a previous cycle of variety development + MAS for SCN and PRR in early generations + phenotypic selection for yield in all field trial stages. 45
Simulation experiment 5: Selection of crosses based on performance of relatives in the previous cycle of variety development + MAS for SCN and PRR in early generation + genomic selection in the PYT from the same training set used for selection of crosses 49
Comparison of experiments 1 and 2: Impact of MAS in early generations for non-yield traits 53
Comparison of experiments 2 and 3: Impact of using GEBV's to select founding crosses based on performance of founders as the training set 55
Comparison of experiments 3 and 4: Impact of using founding crosses based on performance of relatives from a previous variety development project 56

Comparison of experiments 4 and 5: Impact of using GEBV's to select lines in a PYT, when GEBVs are based on performance of relatives' in a previous variety development project 58

Relative efficacy among the five variety development experiments 60

Relative efficiencies between variety development experiments 1 and 2 61

Relative efficiencies between variety development experiments 2 and 3 62

Relative efficiencies between variety development experiments 3 and 4 62

Relative efficiencies between variety development experiments 4 and 5 63

Relative efficiencies between variety development experiments 1 and 5 64

DISCUSSION 65

Variety development project assumptions 65

Genetic architecture 69

Effectiveness metrics 70

Cost effectiveness 71

REFERENCES 73

APPENDIX A1. LIST OF 112 PUBLIC VARIETIES, INCLUDING 27 KEY GERMPLASM AND 85 COMMERCIAL VARIETIES GROWN FROM 1995 TO 2008. 80

APPENDIX A2. SOYBEAN CHROMOSOMES AND LINKAGE GROUPS (LG), WITH THEIR RESPECTIVE SIZE IN CM AND NUMBER OF LOCI CHROMOSOME AND NUMBER OF YIELD LOCI PER CHROMOSOME……..……83

APPENDIX A3. ANALYSIS OF VARIANCE FOR SIMULATED GENOTYPIC VALUES OF YIELD DUE TO FOUNDER SETS, STAGE OF DEVELOPMENT, GENETIC ARCHITECTURES AND HERITABILITY IN EXPERIMENT 1: PHENOTYPIC SELECTION FOR ALL TRAITS……84

APPENDIX A4. SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT OF SET 1 FROM EXPERIMENT 1……...85

APPENDIX A5. SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT OF SET 2 FROM EXPERIMENT 1……...86

APPENDIX A6. SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT OF SET 3 FROM EXPERIMENT 1……...87

APPENDIX A7. LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 1 (SET 1), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES…….....……88

APPENDIX A8. LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 1 (SET 2), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES..... 89

APPENDIX A9. LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 1 (SET 3), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES……..90

APPENDIX A10. DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY IN FOUNDER SET 1 USED IN EXPERIMENT 1 FOR NINE COMBINATIONS OF GENETIC ARCHITECTURE AND HERITABILITY FOR EACH OF FOUR STAGES IN SOYBEAN VARIETY DEVELOPMENT……...91

APPENDIX A11. DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY IN FOUNDER SET 2 USED IN EXPERIMENT 1 FOR NINE COMBINATIONS OF GENETIC ARCHITECTURE AND HERITABILITY FOR EACH OF FOUR STAGES IN SOYBEAN VARIETY DEVELOPMENT……...92

APPENDIX A12. DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY IN FOUNDER SET 3 USED IN EXPERIMENT 1 FOR NINE COMBINATIONS OF GENETIC ARCHITECTURE AND HERITABILITY FOR EACH OF FOUR STAGES IN SOYBEAN VARIETY DEVELOPMENT……...93

APPENDIX A13. ANALYSIS OF VARIANCE FOR SIMULATED GENOTYPIC VALUES FOR YIELD DUE TO STAGE OF DEVELOPMENT, FOUNDER SETS, GENETIC ARCHITECTURES AND HERITABILITY OF LINES FOR EXPERIMENT 2: MAS FOR SCN AND PRR TRAITS + PHENOTYPIC SELECTION FOR YIELD…… 94

APPENDIX A14. SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FOR SET 1 FROM EXPERIMENT 2……95

APPENDIX A15. SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FOR SET 2 FROM EXPERIMENT 2……96

APPENDIX A16. SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FOR SET 3 FROM EXPERIMENT 2……97

APPENDIX A17. LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 2 (SET 1), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES……..98

APPENDIX A18. LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 2 (SET 2), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES…… 99

APPENDIX A19. LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 2 (SET 3), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES……100

APPENDIX A20. DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY IN FOUNDER SET 1 USED IN EXPERIMENT 2 FOR NINE COMBINATIONS OF GENETIC ARCHITECTURE AND HERITABILITY FOR EACH OF FOUR STAGES IN SOYBEAN VARIETY DEVELOPMENT…….101

APPENDIX A21. DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY IN FOUNDER SET 2 USED IN EXPERIMENT 2 FOR NINE COMBINATIONS OF GENETIC ARCHITECTURE AND HERITABILITY FOR EACH OF FOUR STAGES IN SOYBEAN VARIETY DEVELOPMENT…….102

APPENDIX A22. DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY IN FOUNDER SET 3 USED IN EXPERIMENT 2 FOR NINE COMBINATIONS OF GENETIC ARCHITECTURE AND HERITABILITY FOR EACH OF FOUR STAGES IN SOYBEAN VARIETY DEVELOPMENT…….103

APPENDIX A23. ANALYSIS OF VARIANCE FOR SIMULATED GENOTYPIC VALUES FOR YIELD DUE TO FOUNDER SETS, GENETIC ARCHITECTURES AND HERITABILITY OF EYT LINES IN EXPERIMENT 3, IN WHICH FOUNDER CROSSES WERE SELECTED BASED ON TRAINING SETS CONSISTING OF 1,000 SNP ASSAYS AND SIMULATED YIELDS IN 112 FOUNDERS PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD…….104

APPENDIX A24. SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FROM EXPERIMENT 3 FOR YA SET……..105

APPENDIX A25. SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FROM EXPERIMENT 3 FOR YB SET……...106

APPENDIX A26. SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FROM EXPERIMENT 3 FOR YC SET……...107

APPENDIX A27. LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 3 (YA SET), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES……108

APPENDIX A28. LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 3 (YB SET), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES……109

APPENDIX A29. LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 3 (YC SET), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES……110

APPENDIX A30. DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY METRICS FOR YA SET OF FOUNDER CROSSES FOR EACH HERITABILITY GENETIC ARCHITECTURE COMBINATION AND STAGE OF DEVELOPMENT IN SOYBEAN VARIETY DEVELOPMENT. EXPERIMENT 3 CONSISTS OF FOUNDER CROSSES THAT WERE SELECTED BASED ON TRAINING SETS CONSISTING OF 1,000 SNP ASSAYS AND SIMULATED YIELDS IN 112 FOUNDERS PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD……111

APPENDIX A31. DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY METRICS FOR YB SET OF FOUNDER CROSSES FOR EACH HERITABILITY GENETIC ARCHITECTURE COMBINATION AND STAGE OF DEVELOPMENT IN SOYBEAN VARIETY DEVELOPMENT. EXPERIMENT 3 CONSISTS OF FOUNDER CROSSES THAT WERE SELECTED BASED ON TRAINING SETS CONSISTING OF 1,000 SNP ASSAYS AND SIMULATED YIELDS IN 112 FOUNDERS PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD……112

APPENDIX A32. DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY METRICS FOR YC SET OF FOUNDER CROSSES FOR EACH HERITABILITY GENETIC ARCHITECTURE COMBINATION AND STAGE OF DEVELOPMENT IN SOYBEAN VARIETY DEVELOPMENT. EXPERIMENT 3 CONSISTS OF FOUNDER CROSSES THAT WERE SELECTED BASED ON TRAINING SETS CONSISTING OF 1,000 SNP ASSAYS AND SIMULATED YIELDS IN 112 FOUNDERS PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD……113

APPENDIX A33. ANALYSIS OF VARIANCE FOR SIMULATED GENOTYPIC VALUES FOR YIELD DUE TO STAGE OF DEVELOPMENT, FOUNDER SETS, GENETIC ARCHITECTURES AND HERITABILITY OF LINES FOR SIMULATED EXPERIMENT 4 IN WHICH CROSSES CONSISTS OF RELATIVE PREDICTIONS FROM LINES SELECTED FROM AYTS AND EYT IN EXPERIMENT 3, PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD……..114

APPENDIX A34. SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FOR EXPERIMENT 4 (GA SET)……115

APPENDIX A35. SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FOR EXPERIMENT 4 (GB SET)……116

APPENDIX A36. SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FOR EXPERIMENT 4 (GC SET)……117

APPENDIX A37. LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 4 (GA SET), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES……118

APPENDIX A38. LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 4 (GB SET), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES……119

APPENDIX A39. LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 4 (GC SET), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES……120

APPENDIX A40. DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY METRICS FOR GA FOUNDING SETS OF CROSSES USED TO INITIATE VARIETY DEVELOPMENT FOR EACH HERITABILITY AND GENETIC ARCHITECTURE COMBINATION AND STAGE OF DEVELOPMENT FOR EXPERIMENT 4. THE THREE SETS OF CROSSES WERE SELECTED BASED ON PREDICTED PROGENY VALUES, WHERE THE TRAINING SET CONSISTED OF RELATED LINES SELECTED FROM AYTS AND EYT IN EXPERIMENT 3, PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD……121

APPENDIX A41. DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY METRICS FOR GB FOUNDING SETS OF CROSSES USED TO INITIATE VARIETY DEVELOPMENT FOR EACH HERITABILITY AND GENETIC ARCHITECTURE COMBINATION AND STAGE OF DEVELOPMENT FOR EXPERIMENT 4. THE THREE SETS OF CROSSES WERE SELECTED BASED ON PREDICTED PROGENY VALUES, WHERE THE TRAINING SET CONSISTED OF RELATED LINES SELECTED FROM AYTS AND EYT IN EXPERIMENT 3, PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD……122

APPENDIX A42. DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY METRICS FOR GC FOUNDING SETS OF CROSSES USED TO INITIATE VARIETY DEVELOPMENT FOR EACH HERITABILITY AND GENETIC ARCHITECTURE COMBINATION AND STAGE OF DEVELOPMENT FOR EXPERIMENT 4. THE THREE SETS OF CROSSES WERE SELECTED BASED ON PREDICTED PROGENY VALUES, WHERE THE TRAINING SET CONSISTED OF RELATED LINES SELECTED FROM AYTS AND EYT IN EXPERIMENT 3, PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD……123

APPENDIX A43. ANALYSIS OF VARIANCE OF AVERAGE GENOTYPIC VALUES FOR YIELD SHOWING SIGNIFICANCE FOR THE INTERACTION BETWEEN FOUNDER SET, GENETIC ARCHITECTURE AND HERITABILITY, AS WELL STAGE OF DEVELOPMENT, GENETIC ARCHITECTURE AND HERITABILITY. EXPERIMENT 5 CONSISTS OF RELATIVE PREDICTIONS FROM LINES SELECTED FROM AYTS AND EYT IN EXPERIMENT 3, PLUS MAS FOR SCN AND PRR TRAITS PLUS GENOTYPIC SELECTION FOR YIELD USING GEBVS FOR PYT LINES……124

APPENDIX A44. SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FOR EXPERIMENT 5 (GA SET)……125

APPENDIX A45. SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FOR EXPERIMENT 5 (GB SET)……126

APPENDIX A46. SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FOR EXPERIMENT 5 (GC SET)……127

APPENDIX A47. LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 5 (GA SET), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES……128

APPENDIX A48. LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 5 (GB SET), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES……129

APPENDIX A49. LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 5 (GC SET), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES……130

APPENDIX A50. AVERAGE DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY FOR GA FOUNDING SETS OF CROSSES USED TO INITIATE VARIETY DEVELOPMENT FOR NINE COMBINATIONS OF GENETIC ARCHITECTURE AND HERITABILITY AND STAGE OF DEVELOPMENT IN SOYBEAN VARIETY DEVELOPMENT FOR EXPERIMENT 5 IN WHICH CROSSES CONSISTS OF RELATIVE PREDICTIONS FROM LINES SELECTED FROM AYTS AND EYT IN EXPERIMENT 3, PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD……..131

APPENDIX A51. AVERAGE DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY FOR GB FOUNDING SETS OF CROSSES USED TO INITIATE VARIETY DEVELOPMENT FOR NINE COMBINATIONS OF GENETIC ARCHITECTURE AND HERITABILITY AND STAGE OF DEVELOPMENT IN SOYBEAN VARIETY DEVELOPMENT FOR EXPERIMENT 5 IN WHICH CROSSES CONSISTS OF RELATIVE PREDICTIONS FROM LINES SELECTED FROM AYTS AND EYT IN EXPERIMENT 3, PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD……..132

APPENDIX A52. AVERAGE DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY FOR GC FOUNDING SETS OF CROSSES USED TO INITIATE VARIETY DEVELOPMENT FOR NINE COMBINATIONS OF GENETIC ARCHITECTURE AND HERITABILITY AND STAGE OF DEVELOPMENT IN SOYBEAN VARIETY DEVELOPMENT FOR EXPERIMENT 5 IN WHICH CROSSES CONSISTS OF RELATIVE PREDICTIONS FROM LINES SELECTED FROM AYTS AND EYT IN EXPERIMENT 3, PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD……..133

LIST OF FIGURES

Page

Figure 1: Global soybean hectares planted from 1994-2014. 1

Figure 2: Depiction of steps in a soybean variety development project. 5

Figure 3: Soybean diseases and agronomic traits. A) SCN score 1 (left) and 9 (right), B) PRR score 1 (left) and 9 (right), C) Lodging score 1 (left) and 9 (right), and D) aerial view of varieties with different RM planted in strip plots. 6

Figure 4: Stages in which MAS (blue triangle) and GS (black triangle) have been integrated into soybean variety development projects. 13

Figure 5: Stages in which MAS (blue triangle) and GS (gray, black, green triangles) have been integrated into soybean variety development projects: PS (phenotypic selection), MAS + PS (MAS + phenotypic selection), FP + MAS + PS (founder predictions + MAS + phenotypic selection), RP + MAS + PS (relative predictions + MAS + phenotypic selection), RP + MAS + GS (relative predictions + MAS + genomic selection). 22

Figure 6: Hierarchical cluster showing genetic similarity among soybean lines RM I – VII. 30

Figure 7: Distributions of simulated A) genotypic values of yield, B) phenotypic values of yield, C) lodging, D) PRR, and E) SCN of lines evaluated from experiment 1. Distributions are composed of data aggregated from lines across all stages of development and all combinations of simulated genetic architectures and heritabilities. 33

Figure 8: Distributions of simulated A) genotypic values of yield, B) phenotypic values of yield, C) lodging, D) PRR, and E) SCN of lines evaluated from experiment 2. Distributions are composed of data aggregated from lines selected for all combinations of simulated genetic architectures and heritabilities. 37

Figure 9: Distributions of simulated A) genotypic values of yield, B) phenotypic values of yield, C) lodging, D) PRR, and E) SCN of lines evaluated from experiment 3. Distributions are composed of data aggregated from lines selected for all combinations of simulated genetic architectures and heritabilities. 42

Figure 10: Distributions of simulated A) genotypic values of yield, B) phenotypic values of yield, C) lodging, D) PRR, and E) SCN of lines evaluated from experiment 4. Distributions are composed of data aggregated from lines selected for all combinations of simulated genetic architectures and heritabilities. 46

Figure 11: Accuracy (left axis) measured as Pearson`s correlation. The genetic architecture values correspond to 20%, 60%, and 100% of the phenotypic variance on an entry mean basis. Training sets consisted of simulated genotypic and phenotypic data from 1,540 lines evaluated in AYT-1 and 300 lines evaluated in AYT-2 in combination with genotypic assays at 1,000 loci. A) First set of founder lines, B) Second set of founder lines, C) Third set of founder lines. 49

LIST OF TABLES

Page

Table 1: A generic outline of soybean variety development (Gaynor et al. 2017; Hickey et al. 2017). 4

Table 2: Allocation of resources to a typical soybean line development project. 7

Table 3: Proportion and number of crosses among founders of a variety development project listed by RM. 16

Table 4: Allocation of resources in a simulated soybean variety development project. 17

Table 5: Simulated genetic architectures and contributions to phenotypic variability for five soybean traits. 20

Table 6: Example of a confusion table to determine decision accuracy, sensitivity and specificity. 27

Table 7: Resources allocation in terms of plots and US dollars for five experiments. 28

Table 8: Example of genotypic data from 112 founder lines. 30

Table 9: Membership in five clusters of 112 soybean accessions and their respective RMs. 31

Table 10: Representation of founder phenotype file for yield, RM, lodging, SCN and PRR traits. 31

Table 11: Analysis of variance for simulated genotypic values of yield due to founder sets, genetic architectures and heritability for lines in the final stage (EYT) of variety development in experiment 1 where phenotypic selection was conducted on all traits. 32

Table 12: Genotypic values for yield and least square means for simulated yield, lodging, SCN and PRR in experiment 1. Values are listed by RM group and varietal development stages. 34

Table 13: Average decision accuracy, sensitivity, and specificity from three founder sets used in experiment 1 for nine combinations of genetic architecture and heritability for each of four stages in soybean variety development. 35

Table 14: Analysis of variance for simulated genotypic values for yield in the final stage (EYT) of variety development due to founder sets, genetic architectures and heritability of EYT lines for experiment 2 where MAS for SCN and PRR traits was conducted prior to the PYT and phenotypic selection for yield was conducted through all stages of development. 36

Table 15: Genotypic values for yield and least square means for simulated yield, lodging, SCN and PRR in experiment 2. Values are listed by RM group and varietal development stages. 38

Table 16: Average decision accuracy, sensitivity, and specificity from three founder sets used in experiment 2 for nine combinations of genetic architecture and heritability for each of four stages in soybean variety development. 39

Table 17: Number of selected crosses that were consistent among three sets of 300 selected crosses. Selected sets of crosses were based on predicted GEBVs of 200 simulated recombinant inbred lines derived from 1,100 possible crosses that could be used to initiate a variety selection program. 40

Table 18: Analysis of variance for simulated genotypic values for yield due to founder sets, genetic architectures and heritability of lines in experiment 3, in which founder crosses were selected based on training sets consisting of 1,000 SNP assays and simulated yields in 112 founders plus MAS for SCN and PRR in early generations and phenotypic selection for yield in subsequent stages of field trials. 41

Table 19: Genotypic values for yield and least square means for simulated yield, lodging, SCN and PRR in experiment 3, listed by RM group and varietal development stages in which founder crosses were selected based on training sets consisting of 1,000 SNP assays and simulated yields in 112 founders plus MAS for SCN and PRR traits plus phenotypic selection for yield. 43

Table 20: Average decision accuracy, sensitivity, and specificity metrics from three sets of founder crosses for each heritability genetic architecture combination and stage of development in soybean variety development. Experiment 3 consists of founder crosses that were selected based on training sets consisting of 1,000 SNP assays and simulated yields in 112 founders plus MAS for SCN and PRR traits plus phenotypic selection for yield. 44

Table 21: Analysis of variance for simulated genotypic values for yield due to founder sets, genetic architectures and heritability of EYT lines for a simulated experiment in which crosses consists of relative predictions from lines selected from AYTs and EYT in experiment 3, plus MAS for SCN and PRR traits plus phenotypic selection for yield. 45

Table 22: Genotypic values for yield and least square means for simulated yield, lodging, SCN and PRR in experiment 4. Values are listed by RM group and varietal development stages in a simulated experiment in which crosses consists of relative predictions from lines selected from AYTs and EYT in experiment 3, plus MAS for SCN and PRR traits plus phenotypic selection for yield. 47

Table 23: Average decision accuracy, sensitivity, and specificity metrics from three founding sets of crosses used to initiate variety development for each heritability and genetic architecture combination and stage of development for experiment 4. The three sets of crosses were selected based on predicted progeny values, where the training set consisted of related lines selected from AYTs and EYT in experiment 3, plus MAS for SCN and PRR traits plus phenotypic selection for yield. 48

Table 24: Analysis of variance of average genotypic values for yield for EYT lines showing significant interaction between founder set, genetic architecture and heritability. Experiment 5 consists of relative predictions from lines selected from AYTs and EYT in experiment 3, plus MAS for SCN and PRR traits plus genotypic selection for yield using GEBVs for PYT lines..... 50

Table 25: Genotypic values and least square means for simulated yield, lodging, SCN and PRR in experiment 5. Values are listed by RM group and varietal development stages in a simulated experiment that consists of relative predictions from lines selected from AYTs and EYT in experiment 3, plus MAS for SCN and PRR traits plus genotypic selection for yield using GEBVs for PYT lines. 51

Table 26: Average decision accuracy, sensitivity, and specificity from three founding sets of crosses used to initiate variety development for nine combinations of genetic architecture and heritability and stage of development in soybean variety development for experiment 5 in which crosses consists of relative predictions from lines selected from AYTs and EYT in experiment 3, plus MAS for SCN and PRR traits plus phenotypic selection for yield. 52

Table 27: Combined analysis of variance for EYT simulated genotypic values for yield due to founder sets, genetic architectures and heritability of lines for experiment 1 with only phenotypic selection for yield, and experiment 2 simulated with MAS for SCN and PRR traits + phenotypic selection for yield. 53

Table 28: Least Square Means for simulated genotypic values of yield, phenotypic yield, lodging, SCN and PRR. Comparisons are conducted by RM group and varietal development stages for experiments 1 and 2..... 54

Table 29: Combined analysis of variance for EYT simulated genotypic values for yield due to founder sets, genetic architectures and heritability of lines for experiment 2 and 3 simulated with MAS for SCN and PRR traits + phenotypic selection for yield. 55

Table 30: Least Square Means for simulated genotypic values of yield, phenotypic yield, lodging, SCN and PRR. Comparisons are conducted by RM group and varietal development stages for experiments 2 and 3..... 56

Table 31: Combined analysis of variance for EYT simulated genotypic values for yield due to founder sets, genetic architectures and heritability of lines for experiment 3 and 4 simulated with founder and relative prediction plus MAS for SCN and PRR traits plus phenotypic selection for yield. 57

Table 32: Least Square Means for simulated genotypic values of yield, phenotypic yield, lodging, SCN and PRR. Comparisons are conducted by RM group and varietal development stages for experiments 3 and 4. 58

Table 33: Combined analysis of variance for EYT simulated genotypic values for yield due to founder sets, genetic architectures and heritability of lines for experiment 3 and 4 simulated with relative predictions plus MAS for SCN and PRR traits plus phenotypic and GEBV selection for yield. 59

Table 34: Least Square Means for simulated genotypic values of yield, phenotypic yield, lodging, SCN and PRR. Comparisons are conducted by RM group and varietal development stages for experiments 4 and 5. 60

Table 35: Average number of lines selected in EYT for each genetic architecture if the best 10 highest yield at PYT would have been selected for each experiment. 61

Table 36: Resources allocated by task and generations are represented by US dollars and number of plots for experiments 1 and 2. MAS is practiced in F_2 generation for experiment 2. 61

Table 37: Resources allocated by task and generations are represented by US dollars and number of plots for experiments 2 and 3. MAS is practiced in F_2 generation. 62

Table 38: Resources allocated by task and generations are represented by US dollars and number of plots for experiments 3 and 4. 63

Table 39: Resources allocated by task are represented by US dollars and number of plots for experiments 4 and 5. 63

Table 40: Resources allocated by task are represented by US dollars and number of plots for experiments 1 and 5. GS is practiced in PYT for experiment 5 at $4.65/genotyping. 64

NOMENCLATURE

AUC	Area Under Curve
AYT	Advanced Yield Trial
bu/ac	bushel per acre
EYT	Elite Yield Trial
GEBVs	Genomic Estimated Breeding Values
GS	Genomic Selection
MAS	Marker Assisted Selection
MZ	Maturity Zone
PRR	Phytophthora Root Rot
PYT	Preliminary Yield Trial
QTL	Quantitative Trait Locus/Loci
RM	Relative Maturity
ROC	Receiver Operator Curve
SCN	Soybean Cyst Nematode

ABSTRACT

Marker technologies have allowed soybean breeding to exploit genotypic information for traits of various genetic architectures. However, the incorporation of technologies have had qualitative impacts on resources used in variety development projects. The objective of this study was to investigate the impact of integrating Marker Assisted Selection (MAS) for single gene and oligo-genic traits and Genomic Selection (GS) for yield in soybean variety development projects through use of simulations, decision classifier metrics, and cost analysis. The breeding goals of the project are to maximize yield of soybean varieties adapted to maturity zones (MZs) II, III and IV, while assuring that the varieties will not lodge and are resistant to Phytophthora Root Rot (PRR) and one race of Soybean Cyst Nematode (SCN). These goals need to be met while minimizing costs. Results show that MAS for PRR and SCN can be implemented with similar efficacy and greater efficiency than traditional phenotypic selection systems. Integration of GS into variety development projects can be as effective as phenotypic selection for yield, but is not as efficient unless the costs of marker assays are less than $4.65 per sample (line).

INTRODUCTION

Soybean (*Glycine max* (L.) Merrill) is the world's leading source of vegetable oil and plant-based protein. World production for 2014 is estimated at 306.5 million metric tons produced on 117.5 million hectares (Figure 1). North American and South American countries are responsible for 71% of total world´s soybean acreage and 81% of the world's production (FAO, 2017). Among oilseeds, soybeans provide 29% of the world's vegetable oil and 70% of vegetable protein production (USDA, 2017). In the last 20 years the world's soybean production has more than doubled.

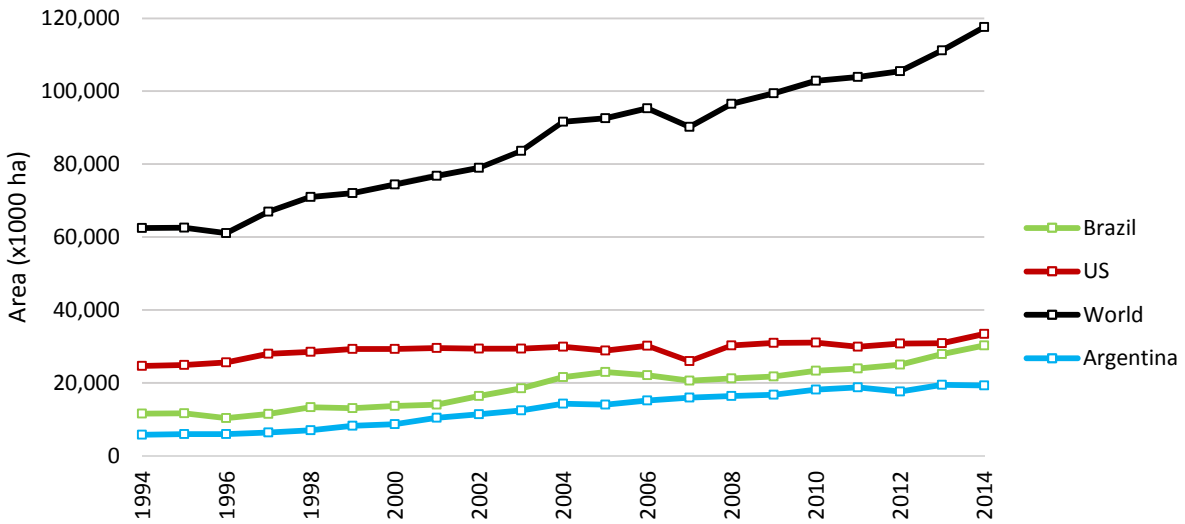

Figure 1: Global soybean hectares planted from 1994-2014.

Annual increases in soybean yields have been estimated to be 23.4 kg/ha/year from 1924 to 2012 for the US and 43.5 kg/ha/year from 1977 to 2013 for Brazil (Specht et al. 2014). The genetic contributions to these changes depend on the heritability and correlations that exist among selected traits. Estimates of heritability for yield have been estimated to be 0.12 to 0.50 (Anand and Torrie, 1963) and 0.83 to 0.89 (Orf et al. 1999). Estimates of heritabilities for lodging and maturity were 0.24 to 0.61 and 0.51 to 0.86 respectively (Anand and Torrie, 1963).

Estimated genetic correlations between yield and relative maturity (RM) are 0.19 to 0.36, yield and lodging are -0.19 to 0.03 and RM and lodging are -0.03 to -0.34, (Johnson et al. 1955; Morrison et al. 2000; Recker et al. 2014).

Estimates of heritability are used to predict genetic gain in a recurrent selection process as the rate of change in the population mean per generation under selection (Falconer and Mackay 1996) and is often expressed as

$$\textbf{Genetic gain} = \frac{\boldsymbol{h^2 . i . \sigma_P}}{\boldsymbol{L}}$$

also known as the breeder's equation where: h^2 is the heritability, i is the selection intensity, σ_p stands for phenotypic standard deviation of the selected trait, and L is the time or years between breeding generations. While the breeder's equation has been used by plant breeders for "back of the envelope" planning, a calculation of Realized Genetic Gain (RGG) is needed to assess actual impacts of plant breeding on increased yields.

At the molecular level, variability among soybean genomes enables heritable responses to selection. Soybean has a genome size of around 1,115-Mb (Schmutz et al. 2009) and encompasses 2,291 cM per meiosis within linkage groups (http://soybase.org/). The soybean genome is an ancient polyploid (paleopolyploid) with at least 2 rounds of duplication and diploidization (Walling et al. 2006). As described by Shultz et al. (2006), the genome is the product of a diploid ancestor (n = 11), which underwent aneuploid loss (n = 10), polyploidization and diploidization.

Soybean is classified as a short-day crop which means that exposure to long day photoperiods prevents most soybean accessions from transitioning from vegetative to reproductive growth phases when grown in high latitude environments. For this reason, soybean accessions are classified as belonging to specific maturity groups (MG). Alliprandini et al.

(2009) suggested that soybean accessions can be classified into 13 maturity groups. Each MG is associated with a geographic region, referred to as a Maturity Zone starting with “000” (triple-zero) for the most northerly zones in Canada and “X” for tropical regions. Each MG is further divided into ten sub-groups that are used to assign a Relative Maturity (RM) value to all accessions, lines and varieties. The RM values are usually calculated based on the number of days from planting until physiological maturity relative to the number of days for check varieties grown in the same field trials within a MZ.

Every year, several soybean varieties are released to market segments defined by their geographic locations. For varieties with published pedigrees it has been noted that all are descendants of a limited number of founders. Mikel et al. (2010) identified 38 cultivars as key varieties used in the pedigrees of most US varieties released from 1970 to 2008. For example, consider the genetic contribution of variety, A3127. In 117 of 494 North American soybean varieties released from 1999 to 2008, the coefficient of parentage of all involves at least 25% of A3127. The excessive use of a few varieties means that genetic diversity in soybean is limited. Mohammadi et al. (2015) remind us that crossing elite lines by elite lines may produce progeny with high average values, but will likely decrease useful genetic variance. In order to address the loss of useful genetic diversity, plant breeders introgress favorable alleles from exotic germplasm (Thompson and Nelson, 1998), from transgenic sources, and cross elite lines with elite lines from other MGs.

In the abstract, soybean variety development consists of four recurring activities: selection of crosses, cross pollination, self-pollination and selection of the best segregating progeny (Table 1). After artificial cross pollination of two homozygous lines, “Line A” x “Line B”, plants are self-pollinated for several generations (F_1 - F_4) before evaluating the segregating

mostly homozygous progeny in field trials at multiple locations for several years. If any of the evaluated lines meet the criteria defined by the breeding objectives, a new soybean variety is released.

Table 1: A generic outline of soybean variety development (Gaynor et al. 2017; Hickey et al. 2017).

Generation	Activity
cross	Homozygous Line A x Homozygous Line B
↓	
$F_1 - F_4$	Self-pollination
↓	
F_4	Select individual plants and use self-pollinated seed to create a line
↓	
$F_{4:5}$	Evaluate lines in an unreplicated field trial, select lines use their self-pollinated seed for future evaluations
↓	
$F_{4:6} - F_{4:9}$	Evaluate selected replicable lines at multiple locations select lines use their self-pollinated seed for future evaluations
Variety	New variety released for commercialization

Variety development in annual crops typically requires 8 – 10 years (Cardinal, 2012; Minella et al. 2015; Carter et al. 2015) and is often depicted as a pipeline (Figure 2). Prior to development of molecular genetic markers, soybean variety development projects relied exclusively on phenotypic information for multiple stages of selection. Visual selection was conducted on individual plants in early generations for highly heritable traits such as plant height and maturity as well as opportunistic traits such as disease resistance and lodging. Yield (per unit land area) requires development of replicable genotypes, a.k.a. lines. Because it is desirable for members of lines to be mostly homozygous and homogeneous it takes several generations of self-pollination to create replicable lines. Thus, lines are not evaluated until several years after the initial crosses in a preliminary yield trial (PYT). While individuals within a line are replicable, the number of seeds per line are usually not sufficient for lines to be replicated. Therefore the PYT is not replicated, so the repeatability of line performance for this stage is

unknown. It takes several additional generations to produce sufficient seed to evaluate lines in replicable trials. The advanced yield trials (AYT_n) include replications of the lines across multiple locations and years, with retention of subsets of lines based on performance for evaluation in subsequent stages of development. Eventually a few elite lines are evaluated for yield in dozens to hundreds of locations in elite yield trials (EYT). The EYT's sample as many farm conditions as affordable before the release of the new soybean variety. If a variety becomes widely accepted by farmers, it will become a cultivated variety, a.k.a., a variety.

Figure 2: Depiction of steps in a soybean variety development project.

While high yielding varieties represent the primary objective of soybean breeding programs, variety development also requires selection for multiple traits, some of which are non-negotiable from the perspective of the farmer. Examples of non-negotiable traits include herbicide resistance in North America, and soybean rust resistance in Brazil. In the United States SCN is considered the most economically damaging pathogen of soybean production (Niblack, 2005). Additional important agronomic traits include PRR, lodging, plant height, and maturity (Figure 3). SCN resistance is scored as an ordinal categorical trait: Resistant = 1, Moderate resistance = 5, and Susceptible = 9. PRR is likewise an ordinal categorical trait but has five categories: Resistant = 1, Moderate Resistance = 3, Moderate Susceptible = 5, Susceptible = 7 and Death = 9. Lodging is an ordinal categorical trait with nine categories where erect plants are scored 1 and prostrate plants are scored 9. RM and yield are quantitative traits that can be measured on continuous scales, although both are recorded as discrete units. Maturity is recorded

as days between planting and physiological maturity and converted to RM values (Figure 3). Yield is recorded as bushels per acre or tons per hectare.

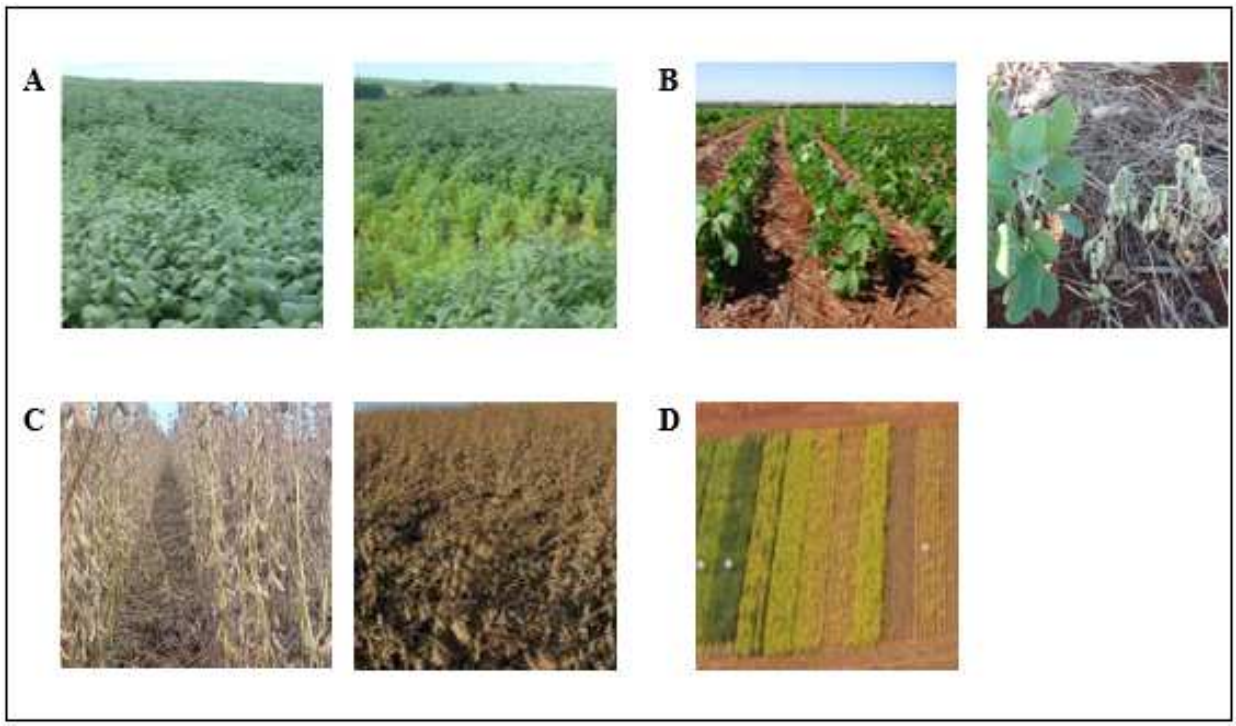

Figure 3: Soybean diseases and agronomic traits. A) SCN score 1 (left) and 9 (right), B) PRR score 1 (left) and 9 (right), C) Lodging score 1 (left) and 9 (right), and D) aerial view of varieties with different RM planted in strip plots.

The theory for simultaneous selection of multiple traits, i.e., selection indices, was developed many decades ago (Smith, 1936; Hazel, 1943), but has not been as successfully implemented in plant breeding as animal breeding. Part of the challenge with using selection indices in plant breeding is that decisions about relative economic weights to place on various traits need to involve individual farmers and their tolerance for risk at the time a breeding project begins, not five to ten years after the project is initiated.

Also, it is good to emphasize that a new variety development project (set of crosses) begins every year in a variety development program. Thus, the soybean breeder has to allocate limited resources to multiple development projects involving all of the steps every year. To

illustrate, consider a typical allocation of resources to each of the developmental stages (Table 2).

Table 2: Allocation of resources to a typical soybean line development project.

Generation	**Unit**	**Unit**	**Plot number**	**Number of location and (rep)**
Possible Crosses	Founders	100 (4,950)	500	1 (5)
Selected Crosses	Families	300		
F_1	Seeds/family	10	300	1 (1)
F_2	Seeds/family	1,000	300	1 (1)
F_3	Seeds/family	1,000	300	1 (1)
F_4	Seeds/family	1,000	300	1 (1)
$F_{4:5}$ - PYT	Lines/family	50 (25-75)	15,000	1 (1)
$F_{4:6}$ - AYT-1	Lines	1,500	7,500	5 (1)
$F_{4:7}$ - AYT-2	Lines	300	7,200	12 (2)
$F_{4:8}$ - EYT	Lines	30	1,800	30 (2)
Variety	Lines	5		

It is not unusual for soybean breeders to begin a variety development project using about 100 lines in a crossing block. All pairwise (without reciprocal) combinations of 100 possible founders could be used to create 4,950 families. However, the budget typically constrains soybean breeders from creating more than about 300 segregating families. While a random set of 300 crosses could be considered to represent the population structure involving the 100 lines, in practice, the 300 crosses are selected. The selected crosses are subsequently used to produce F_1 seeds which are self-pollinated to produce no more than 1,000 F_2 individuals per family. No more than 3,000 F_3 seeds per family, equally representing all F_2 plants, are harvested, but only a sample of 1,000 F_3 seeds are planted and allowed to self-pollinate. No more than 3,000 F_4 seeds per family, equally representing all F_3 progenitors are harvested, and again just a sample of 1,000 F_4 seeds are planted and allowed to self-pollinate. About 50 $F_{4:5}$ lines per family (although some families are represented by as few as 25 or as many as 75 lines) will be selected based on RM

and various opportunistic traits. All F_5 seeds from each selected F_4 plant will be harvested and bulked to create a $F_{4:5}$ line that will be evaluated in PYT. In the PYTs, $F_{4:5}$ lines consisting of about 300 families with 50 lines each, are planted in single rows, evaluated, and harvested. About ten percent of the $F_{4:5}$ lines will be selected. Each line's seed will be bulked to create $F_{4:6}$ lines that can be grown at multiple locations in a total of about 7,500 plots in the first year of replicated AYTs. About twenty percent of the $F_{4:6}$ lines will be selected and $F_{4:7}$ seed will be used in a second year of AYT's at a larger number of locations, resulting in about 7,200 plots. Ten percent of the $F_{4:7}$ lines are selected for EYTs consisting of $F_{4:8}$ lines that will be grown at about 30 locations in 1,800 plots. A successful variety development project will be completed with the decision to select about five to ten varieties for evaluation and product placement by agronomists (sales associates) using on-farm strip trials.

While soybean variety development projects are very similar to that outlined in Table 2, soybean breeders also need to improve the effectiveness and efficiency of the pipeline. Neus (2010) proposed that genetic markers could increase genetic gain for complex traits in soybean by increasing the selection intensity. In addition, molecular genetic markers increase the accuracy of selection by identification of individuals carrying desired simply inherited traits in any generation, even when planted in off-season or out-of-target markets (Collard et al. 2005).

A molecular genetic marker is a DNA variant used to identify differences at genetic loci between individuals (Collard et al. 2005). If these differences between markers are associated with phenotypic differences, they can be used for various types of Marker Assisted Selection (MAS). In the identification phase marker trait associations, a.k.a. Quantitative Trait Loci (QTL) are detected using samples of mostly homozygous lines sampled from populations that exhibit repeatable variability for phenotypic traits of interest (Guo et al. 2013). In the confirmation

phase, genetic markers linked to desirable QTLs are validated and subsequently used as surrogates for selecting phenotypic traits that are expensive or difficult to assay.

MAS for simply inherited and oligo-genic traits has been successfully implemented in varietal development of annual crops because it is a straightforward technique for finding and applying genetic markers at any stage of development (Figure 4). Bonnett et al. (2005) stated that molecular markers allow more accurate selection when used in early generations than conventional methods and Sebastian et al. (2010) demonstrated that MAS in advanced yield trials for development of elite soybean varieties with success. Also MAS has been used in backcrossing and pyramiding of multiple QTL (Werner, et al. 2005). Xu & Crouch (2008) proposed a more general perspective of using MAS for developing breeding strategies for simultaneous improvement of multiple traits.

In commercial soybean varietal development programs, MAS has been used for at least two decades and the number of varieties that have been developed using genetic markers are too numerous to list for purposes of this dissertation. A few illustrative examples include: a) Arelli et al. (2015) released the variety JTN-5203 with resistance to multiple SCN resistance using greenhouse pot screening and six SSR markers to identify and confirm SCN resistance; b) Cianzio et al. (2016) used MAS to develop sudden death syndrome (SDS) and SCN resistant plants and released the variety AR11SDS; c) Diers et al. (2014) registered eight soybean varieties that had been selected for resistance to different soybean rust genes using MAS; d) MAS has been used for gene ‘pyramiding’ transgenic herbicide and insect resistance while avoiding the transfer of undesirable alleles in transgenic lines (Pengyin, 2017).

QTL involved in the expression of more complex polygenic traits are known in several crops, but these are typically family dependent and in most cases the detection is incomplete,

non-replicable and estimates of contributions of genetic variability are biased (Beavis, 1994). Consequently, application of MAS for QTLs associated with traits such as yield have had limited success. In an effort to overcome the bias and validation efforts involved in MAS for quantitative traits, animal breeders extended the “animal model” of mixed linear equations to include information on marker trait associations for all available polymorphic markers (Meuwissen et al. 2001). The method calculates Genomic Estimated Breeding Values (GEBVs) for breeding individuals (or lines) as the sum of all underlying estimated marker trait associations. Their approach has become widely known as Genomic Selection (GS).

In principle, GEBVs can be estimated using phenotypic and genotypic data from any stage of a varietal development program (Figure 2 and Table 2) and likewise applied to segregating progeny in early stages or segregating lines at later stages. Thus, various selection strategies, based on GEBVs, have been proposed for variety development programs (Hickey et al. 2014; Gaynor et al. 2017; He et al. 2016), in oligo-genic traits such as SCN resistance (Bao et al. 2014) or in polygenic traits (He et al. 2016).

The process of obtaining GEBVs involves both a training set consisting of genotypic scores and phenotypic values for traits of interest and a prediction (validation) set consisting of only genotypic scores. The measured phenotypic values are used along with genotypic scores in the training set to obtain estimates of the GEBVs for members of the prediction set. In varietal development projects, phenotypic values for traits of interest are assessed among a set of lines replicated at multiple locations (Moser et al. 2015) and the model is used to predict the GEBVs of a much larger set of lines for which phenotypic performance is not assessed. Experimental evaluation of GEBVs in self-pollinated crops has been reported in wheat (Poland et al. 2012), oats (Asoro et al, 2013), soybean (Jarquín et al. 2014), and rice (Spindel et al. 2015).

At least 14 computational methods have been proposed to obtain GEBVs (Howard et al, 2014). As long as the underlying genetic architecture of a trait is primarily due to additive genetic effects, most of these methods have similar accuracy (Howard et al, 2014; 2017). Accuracy is defined as the correlation between GEBVs and true breeding values which are unknown, except in simulations. In practice, animal breeders define prediction accuracy as the proportion of narrow sense heritability explained by variance of GEBVs (Dekkers, 2007).

Estimated accuracies of GEBVs in validation sets have been used to infer the efficacy of GS. Based on theory, variability among GEBVs should be less than or equal to the narrow sense heritability of the trait (Dekkers, 2007). Thus, high values for estimated prediction accuracies should be expected for highly heritable traits. Moser et al. (2015), obtained estimated prediction accuracies of 0.13 (±0.041), 0.32 (±0.038), and 0.50 (±0.032) for simulated traits in which the heritabilities were 0.2, 0.5, and 0.8 respectively. GS studies conducted in self-pollinated crops report accuracies ranging from 0 to 0.87. Some authors (Lian et al. 2014; Riedelsheimer et al. 2013; Würschum et al. 2013) have reported negative estimates of accuracies for polygenic traits. Ceron-Rojas et al. (2015) reported estimated prediction accuracies that decreased across several successive selection cycles, suggesting that the training sets need to be updated on a regular basis.

Given reduced accuracy of selection relative to phenotypic selection, it is not clear whether the ability to increase selection intensity using GS will justify its adoption in variety development programs. At a more fundamental level, should estimates of Pearson's correlation or Dekker's proportion of heritability be an arbiter for decisions about implementing GS in variety development programs? These are descriptive statistics, much like the coefficient of variation and with sufficient experience, a breeder will begin to have some intuition about

"good" values for the trait of interest. In practice, we need to recognize that GEBVs are going to be used to make decisions. Some individuals or lines will be retained for further evaluations and eventual release as varieties while most will be discarded. In essence, the selection process converts the continuous trait values, including GEBVs, into binary classifier metrics. ***Decision accuracy*** of binary classifier metrics is defined as the proportion of correctly retained and correctly discarded relative to the total number of decisions. Thus, selection accuracy is composed of the proportion of lines that are correctly kept and the proportion of lines that are correctly discarded. The former is defined as the ***decision sensitivity*** and the latter is defined as the ***decision specificity***.

Surprisingly, plant breeders have not used classifier metrics to evaluate proposed changes in methodology or adoptions of new technologies for variety development. This is due, in part, to the extremely expensive and time consuming experiments needed to investigate hypotheses about contributions of new technologies in a variety development program. In the literature we found only one instance in which 13 methods for aggregating partially ranked data from plant breeding trials were evaluated for their impact on decision-making (Simko & Pechenick, 2010). This is surprising because development of new methods and technologies in allied disciplines of molecular and computational biology routinely use these metrics. For example, Yang et al. (2016) assessed the decision accuracy of four methods for assigning genotypic values to images from the AmpSeq platform.

GS and MAS have been implemented in commercial seed organizations in the early stages of variety development (Figure 4). It is not clear whether decisions to include MAS and GS in variety development projects were based on decision classifier metrics, nor have impacts of these changes been reported. Indeed, there are leaders in commercial seed organizations who

have suggested that these decisions were not based on objective and measurable criteria such as costs and decision classifier metrics (J. Byrum, personal communication, 2016).

Figure 4: Stages in which MAS (blue triangle) and GS (black triangle) have been integrated into soybean variety development projects.

Simulations based on accepted models for inheritance and genetic architectures can be used to assess impacts of breeding decisions on genetic improvement (Podlich and Cooper, 1999; Sun et al. 2011; Li et al. 2012). Sun et al. (2011) described available software packages (QUGENE/QuLine, MPB, GREGOR, PLABSIM, GENEFLOW and COGENFITO) used for evaluating various plant breeding programs, including marker-assisted backcross procedures, effect of selection on genetic diversity, effect of genotype-by-environment interactions on minimal resources needed to maximize the genetic gain. Podlich & Cooper (1999) developed QUGENE to compare two selection strategies associated with individual's performance in a single and multiple environments, while accounting for epistasis and genotype x environment interaction (GxE). Wang et al. (2003) used QUGENE and QUCIM to simulate and compare two of CIMMYT's breeding systems in terms of genetic gain, number of crosses and resource allocation. Their simulations demonstrated that a bulk selection method resulted in genetic gain that was 3.3% higher than in the modified pedigree/bulk selection method. Bernardo (2014a) developed his own simulation software and reported the mean and standard deviation of GS on accuracy and selection response from 1,000 simulations of genetic architectures that include alleles with "major effects". Longin et al. (2015) used the open source software package "selectiongain" to identify the optimal combination of number of lines, locations and testers for

different breeding strategies. From among the strategies, they found the greatest annual genetic gains with an early stage of GS followed by phenotypic selection.

Herein, we investigate the impact of integrating MAS for single gene and oligo-genic traits and GS for yield in soybean variety development through use of simulations, decision classifier metrics, and costs. The breeding goals of the variety development project are to maximize yield of varieties adapted to maturity zones II, III and IV, while assuring that the varieties will not lodge and are resistant to PRR and race 1 of SCN. These goals need to be met while minimizing costs.

The objectives of the research reported herein are to assess the relative effectiveness and efficiency of MAS strategies that have been integrated into soybean variety development projects during the last 20 years. Explicitly, we conducted five consecutive simulation experiments that reflect stepwise changes that have been implemented in commercial soybean variety development projects. Each consecutive simulation experiment beginning with the second consists of conditions that exist in the prior experiment plus one modification to the variety development process.

We first simulated replicated soybean variety development projects that use only phenotypic selection. We hypothesize that random subsets of crosses from among crosses with high potential will not affect genetic potential nor sensitivity and specificity in selecting the ten highest yielding varieties. We refer to this as the *phenotypic selection* experiment. Second, we hypothesize that application of MAS for single and oligogenic traits in early stages of variety development will not affect the ability to select the ten best yielding varieties. We utilize the same sets of founder crosses used in the phenotypic selection experiment and apply MAS to non-yield traits (SCN and PRR) in the early stages and refer to this experiment as *MAS + phenotypic*

selection. Subsequently we compared the genetic potential, specificity, sensitivity and costs of *MAS+phenotypic selection* with results from *phenotypic selection* experiment. Third, we hypothesize it is not possible to use genomic prediction for yield based on data from the founders to accurately predict the best set of crosses to initiate a development project. We refer to this third simulation experiment as *Founder predictions+MAS+phenotypic selection* and compare its outcomes, in terms of genetic potential, specificity, sensitivity and costs, to the same outcomes of *MAS+phenotypic selection*. It has been demonstrated that training sets based on small numbers, e.g., 112 founders, are not associated with accurate models. Also, training sets that utilize relatives of the validation sets provide more accurate models. Fourth, we hypothesize that it is not possible to improve variety development outcomes by using genomic prediction for yield based on training sets derived from relatives to select crosses for initiating the development project. We refer to this fourth simulation experiment as *Relative predictions+MAS+phenotypic selection* and compare its outcomes, in terms of genetic potential specificity, sensitivity and costs, to the same outcomes of *Founder predictions+MAS+phenotypic selection*. Last we hypothesize that use of genomic prediction for yield during the early development stages will not alter the outcomes, in terms of genetic potential specificity, sensitivity and costs, in selecting the ten best varieties. We refer to this fifth simulation experiment as *Relative predictions+MAS+genomic selection* and compare its outcomes, in terms of genetic potential specificity, sensitivity and costs, to the same outcomes of *Relative predictions+MAS+phenotypic selection*.

METHODS

Variety development

An assessment of existing simulation software for plant breeding indicated that it was not possible to model nuances of soybean variety development projects depicted in Table 2. For example, because RM has a strong impact on conducting soybean field trials, decisions about which crosses should be selected to initiate a variety development project involve unequal proportions based on RM values of the founders (Table 3). In order to assure that all RM values are adequately represented throughout all stages of development, each family created by a cross of founders needs to be classified into three groups based on differences of 0, 1 or 2 RM values between the pair of lines used as family founders. Later when $F_{4:5}$ lines are derived, stratified random samples of the F_4 plants needed to be simulated to create 25, 50 and 75 $F_{4:5}$ lines per family based on 0, 1, and 2 RM differences. On the average, each family needs to be represented by about 50 lines, but just as importantly the simulated sampling needs to produce lines representing the full range of RMs for the targeted MZs. In the final stages of the process, simulated RM values of F_4 derived lines were categorized as belonging to one of four RM groups: RM 1.8-2.4, 2.4-3.0, 3.0-3.6, and 3.6-4.2 and used for purposes of selection within RM groups.

Table 3: Proportion and number of crosses among founders of a variety development project listed by RM.

RM	Crosses by RM (%)	Crosses by RM
2.0/2.0	10%	30
2.0/3.0	20%	60
2.0/4.0	35%	105
3.0/3.0	10%	30
3.0/4.0	15%	45
4.0/4.0	10%	30
Total	**100%**	**300**

Since limitations of software should not determine which research questions to pursue, we decided to conduct simulations with assistance from Shengchu Wang, a C/C++ software developer. The resulting simulations enabled simulation of population structure, genetic architecture, self-pollination, genetic recombination, selection strategies and activities associated with resource allocations in a soybean variety development project for MZ's II to IV (Table 4).

Table 4: Allocation of resources in a simulated soybean variety development project.

Generation	**Unit**	**Number of Units**	**Number of Plots**	**Number of locations and replicates per location ( )**
Planned Cross	Families	6,216	560	
Selected Cross	Families	300		
F_1	Seeds/family	10	300	
F_2	Seeds/family	3,000	300	
F_3	Seeds/family	3,000	300	
F_4	Seeds/family	300	300	
$F_{4:5}$ - PYT	Lines/family	50 (25-75)	15,375	1 – (1)
$F_{4:6}$ - AYT-1	Lines	1,540	7,700	5 – (1)
$F_{4:7}$ - AYT-2	Lines	300	7,200	12 – (2)
$F_{4:8}$ - EYT	Lines	40	2,400	30 – (2)
Variety	Lines	10		

Genotypic data from potential founders

One hundred and twelve public soybean accessions (Appendix A1) adapted to MZ's I - VII were considered for use in founder crosses for five simulated variety development projects. The soybean accessions included 27 that were major contributors to pedigrees of modern soybean varieties (Mikel et al. 2010) and 85 were commercial varieties grown by farmers from 1995 to 2008.

The genomes of all 112 accessions were represented by genotypic scores from assays based on the SoySNP50k chip and are available at Soybase (http://soybase.org/). SNPs were distributed across 20 chromosomes (Appendix A2) according to previously determined locations

(Song et al. 2016). Markers with minor allele frequencies of less than 1% were discarded resulting in 34,708 SNP loci distributed randomly among all 20 linkage groups of the 112 founders. The proportion of matching SNPs (Nei and Li, 1979) between all pairs of founders were used as a similarity metric for an agglomerative hierarchical clustering (Goodman and Lasker, 1974) and plotted (JMP 12.0.1, SAS Institute) to illustrate the population structure of the founders. Relationships between RMs of the founders and their associations with the population structure were tabulated.

Simulated genomes of progeny in filial generations evaluated at each stage of development

Methods for simulating genetic inheritance of genomes are well established (Fraser and Burnell, 1970) and were adapted for soybean genomes and stages of variety development. Briefly, haploid gametes from each of the founders were digitally represented by 112 vectors of genotypic scores at 34,708 SNP loci. All possible pairs of founder vectors were combined to create 6216 matrices, each consisting of 34,708 x 2 elements with genotypic scores representing diploid genomes of all possible F_1 progeny. For the F_1 and all subsequent filial generations, each column of the matrices was created by randomly pairing vectors representing haploid gametes created by the previous filial generation.

In order to simulate vectors representing a gamete produced in every generation, a vector with estimated genetic recombination between adjacent pairs of SNP loci was based on prior estimates of map positions (cM) in segregating progeny from Williams 82 x *G. soja* PI 479752 (Song et al. 2016). Map positions were converted to recombination units by assuming no interference among cross-over events and Haldane's function, $\boldsymbol{r = 0.5\,[1 - e^{(-2M)}]}$, where M is the distance between adjacent pairs of loci in Morgans. The vector of values for *r* were compared with a random number sampled from a U[0,1] distribution for every pair of adjacent loci, located

at positions (λ and λ+1) in each matrix representing every individual (or line). If the random number was less than ***r***, then column 1 of the 34708 x 2 matrix was replaced with column 2 beginning at position λ+1 of the matrix and included all rows indexed with values greater than λ+1. Additional recombination events, represented as exchanges of subsets of the columns, were simulated in the same manner using all adjacent rows in each matrix representing a diploid individual. One of the two gametes was randomly obtained to pair with another random gamete produced in the same manner from the same individual to produce a selfed progeny.

Simulated phenotypic values

SCN resistance for one race was modeled as consisting of one major genetic locus on chromosome 10 and all of the phenotypic variability was simulated to consist entirely of genotypic variability (Table 5). PRR resistance was modeled as consisting of two QTL on chromosome 1 and 11 that are respectively responsible for 60% and 30% of the total phenotypic variance (Zhang et al. 2017). The genetic architecture for lodging resistance consisted of five QTL, two on chromosome 15, one on chromosome 17 and two on chromosome 19. Allelic variants at each locus accounted for 16% of the total phenotypic variance and non-genetic factors contributed 20% to the phenotypic variability. The genetic architecture for maturity included five QTL on chromosomes 1, 7, 8, 9, and 10. The first two had large genetic effects responsible for 39.26% and 20.98%, of the total phenotypic variance, the remaining three were responsible for 17.74%, 1.19%, and 1.18% of the phenotypic variance and non-genetic sources of variability contributed 20% to the total phenotypic variance.

Table 5: Simulated genetic architectures and contributions to phenotypic variability for five soybean traits.

Trait	Number of loci	Chromosome	H*	Distribution of additive genetic effects
Yield	3,000	1 to 20	0.2, 0.6, 1.0	Equal
Yield	3,000	1 to 20	0.2, 0.6, 1.0	Negative Exponential
Yield	3,000	1 to 20	0.2, 0.6, 1.0	Three distinct magnitudes
RM	5	1, 7, 8, 9, 10	0.8	Normal
Lodging	5	15, 17, 19	0.8	Uniform
SCN	1	10	1.0	Binomial
PRR	2	1, 11	0.9	Multinomial

* Proportion of phenotypic variance on an entry mean basis due to F_4 derived lines.

The genetic architectures for yield (bu/ac) utilized 3,000 SNP loci distributed randomly across the 20 linkage groups (Appendix A2). Simulated yield values for individuals in the early stages of variety development (Figure 2) were modeled as

$$\boldsymbol{y_i = \mu + G_i + e}$$

where $\boldsymbol{y_i}$ is the phenotypic value, $\boldsymbol{\mu}$ is the overall mean, $\boldsymbol{G_i}$ is the additive genetic value of the i^{th} simulated individual (or line) and $\boldsymbol{e}$ represents the non-genetic sources of variability that were sampled from a N(0, σ^2). The additive genetic value was simulated as $\boldsymbol{G_i = \sum_{j=1}^{p} x_{ij}\, b_j}$, where $\boldsymbol{x_{ij}}$ is the genotypic score at the j^{th} locus in the i^{th} individual, and $\boldsymbol{b_j}$ is the allelic substitution effect at the j^{th} marker locus.

Magnitudes of the allelic substitution effects assigned to each locus were sampled from three distributions. In the first, equal additive effects were assigned to the alternate alleles at each of the 3,000 loci. At the other extreme it can be argued that there are many loci with alleles that each contribute small additive effects and a few loci with alleles that contribute large additive genetic effects. We modeled this argument by sampling genetic effects from a negative exponential distribution. As a compromise between the two extreme distributions of genetic effects, we assigned equal additive effects responsible for 15% of the genotypic variance to 10

loci, equal additive effects responsible for 50% of the genotypic variance to 310 loci and equal additive effects responsible for 35% of the genotypic variance to 2,680 loci. For each of the distributions of additive genetic effects, the non-genetic sources of variability contributed 80%, 40% or 0% to the phenotypic variance of yield, resulting in nine simulated genetic architectures.

Simulated phenotypic values for lines evaluated in field trials for various stages of variety development (Figure 2) were modeled as

$$\boldsymbol{y_{ijk} = \mu + L_j + R_{k(j)} + G_i + (GL)_{ij} + e_{ijk}} ,$$

where $\boldsymbol{y_{ijk}}$ is the phenotype of the ith line (also known generically as a genotype) in the k^{th} field plot within the j^{th} location; $\boldsymbol{\mu}$ is the overall mean; $\boldsymbol{L_j}$ is the non-genetic effect of location j; $\boldsymbol{R_{k(j)}}$ is the non-genetic effect of the kth plot within the j^{th} location; $\boldsymbol{G_i}$ is the additive genetic effect of the i^{th} line; $\boldsymbol{(GL)_{ij}}$ is an interaction effect of the j^{th} location and the line, and $\boldsymbol{e_{ijk}}$ represents additional non-genetic sources of variability sampled from a N(0, σ^2) distribution. The genotypic value for the ith line was simulated as $\boldsymbol{G_i = \sum_{l=1}^{p} x_{il}\, b_l}$, where $\boldsymbol{x_{il}}$ is the genotype score at the l^{th} locus of the genome in the i^{th} line, and $\boldsymbol{b_l}$ is the allelic substitution effect at the l^{th} marker locus.

Simulation experiments

We conducted five simulation experiments that mimic consecutive implementations of MAS strategies that have been used in commercial soybean variety development projects. After the first simulation experiment, each consecutive simulation experiment consists of the same conditions that exist in the prior experiment plus one modification to the variety development process (Figure 5).

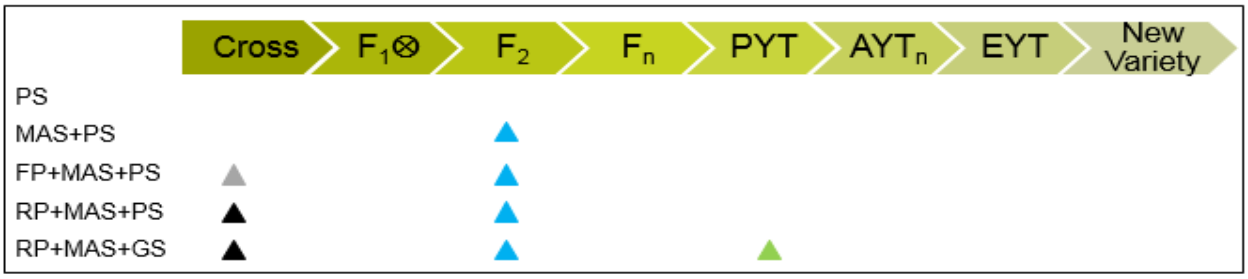

Figure 5: Stages in which MAS (blue triangle) and GS (gray, black, green triangles) have been integrated into soybean variety development projects: PS (phenotypic selection), MAS + PS (MAS + phenotypic selection), FP + MAS + PS (founder predictions + MAS + phenotypic selection), RP + MAS + PS (relative predictions + MAS + phenotypic selection), RP + MAS + GS (relative predictions + MAS + genomic selection).

Simulation experiment 1: Phenotypic selection experiment.

Simulated phenotypic values for all five traits in the set of 112 founders, described above, were used to identify potential crosses that would likely produce undesirable progeny. For example, based on Mendelian inheritance, a potential cross involving two founders that are both susceptible to SCN will produce a family in which all lines are susceptible to SCN. This would be classified as an undesirable cross as would crosses involving susceptibility to PRR and lodging. Also, since the lines need to be adapted to maturity zones II, III and IV, a pair of founders with RM values that are greater than 4.0 or less than 2.0 would not be desirable. Based on these criteria, only 1,100 of the 6,216 possible pairs were considered desirable crosses.

In this first experiment there was no additional information upon which to choose 300 crosses from among 1,100. Because useful genetic variability of the initial crosses could have a significant impact on outcomes, we replicated the first variety development experiment with three random samples of 300 crosses (Table 4). Subsequent samples of segregating individuals and lines used a process of stratified random sampling, as previously described (Table 2) to assure that RM groups were represented by the same numbers of F_4 derived lines in the PYT, AYT and EYT stages (Table 4). Selection was not practiced from the F_1 to F_3 generations (Table

4). In the generation immediately prior to the PYT, F_4 individuals with simulated phenotypic lodging scores less than "6" were selected. Among the selected F_4 individuals, a stratified random sample of 25, 50 and 75 $F_{4:5}$ lines per family with 0, 1, and 2 RM differences between their founders were obtained. As a result, 15,375 $F_{4:5}$ lines were evaluated for all non-yield traits and nine simulated yield genetic architectures. Based on simulated yields for each genetic architecture, $F_{4:5}$ lines were evaluated for yield at a single location in the PYT. Ten percent were selected to be $F_{4:6}$ lines for evaluation in the first AYT (Table 4). Simulated yields for each genetic architecture, averaged across all locations for each $F_{4:6}$ line were compared with the average simulated yields for the RM group to which the line was a member. Seventy-five $F_{4:6}$ lines with the highest simulated phenotypic yields (for each architecture) were selected for each of the four RM groups. *In silico*, these were used to produce $F_{4:7}$ lines for the second multi-location AYT. In the second AYT simulated $F_{4:7}$ lines that were SCN resistant, PRR resistant and with lodging scores less than a value of "6" were selected. Among the selected $F_{4:7}$ lines simulated phenotypic yields averaged across all locations for each line were compared with the average simulated phenotypic yields for the RM group to which the line was a member and ten $F_{4:7}$ lines with the highest average simulated phenotypic yield values were selected from each RM group. In aggregate, there were 40 selected $F_{4:7}$ lines, ten from each RM group. These were advanced to $F_{4:8}$ that were evaluated in an EYT across 30 locations that spanned the four MZ's. From among the 40 $F_{4:8}$ lines, 10 were selected based on SCN resistance, PRR resistance lodging scores 1 - 3, and the highest averaged simulated phenotypic yield values across all four RM groups.

Simulation Experiment 2: MAS + phenotypic selection experiment.

This second simulation experiment utilized the same three random sets of crosses used in the *phenotypic selection* experiment. However, rather than waiting until the F_4 generation to select for the non-yield traits, alleles for molecular markers that had previously been mapped to regions in close proximity to alleles associated with the desirable phenotypes were used to select individual plants in the F_2 generations derived from each set of crosses. After MAS, subsequent selection for yield, as determined by all nine genetic architecture models, was conducted using the same process used for PYT, AYT-1, AYT-2 and EYT stages as used in the *phenotypic selection* experiment.

Simulation experiment 3: Founder predictions +MAS+phenotypic selection.

The third simulation experiment utilized a training model based on simulated genotypic and phenotypic yield data for each of the genetic architectures applied to the 112 potential founders. The SNP genotypes at 1,000 loci were used in combination with the simulated yield values for the founders to predict Genomic Estimated Breeding Values (GEBVs) of 200 recombinant inbred lines from all possible 1,100 desirable crosses using the *PopVar* R package (Mohammadi et al, 2015; R Development Core Team, 2016). For each simulated genetic architecture, 300 of 1,100 predicted distributions with the greatest utility values were used as sets (designated YA, YB, and YC) to initiate the variety development project. Subsequent development of ten varieties from each set followed the same process as that of the second experiment.

Simulation experiment 4: Relative predictions +MAS+phenotypic selection.

Rather than using the genotypes and simulated phenotypes from the 112 founders to train a model for selecting sets of 300 crosses to initiate variety development, we utilized performance

of lines evaluated a previous cycle of variety development to train a model. To approximate the use of data from a previous cycle of variety development that will be used as a training set to calculate GEBVs, the genotypes and phenotypes of selected lines evaluated in the AYT trials of experiment 3 were used as training sets. In other words, training sets consisted of simulated genotypic and phenotypic data from 1,540 lines evaluated in AYT-1 and 300 lines evaluated in AYT-2 for each of the nine genetic architectures responsible for yield in each of the three sets of families. Explicitly we used SNP genotypes at 1,000 loci in combination with the simulated yield values and Ridge regression best linear unbiased prediction (RR-BLUP) implemented in R (Endelman, 2011) to obtain GEBVs:

$$\mathbf{z} = \boldsymbol{\mu} + \sum \boldsymbol{\beta}_i \mathbf{X}_i + \mathbf{e}$$

where z is the BLUP of GEBV; $\boldsymbol{\mu}$ is the overall mean; $\boldsymbol{\beta i}$ is effect of marker i treated as random effect with $\boldsymbol{\beta}_i \sim N(0, \sigma^2_g)$; $\boldsymbol{X_i}$ is genotype of SNPs coded as 1(homozygote), 0 (heterozygote), -1 (other homozygote); $\boldsymbol{e}$ is residual effect and assumed as $N(0, \sigma^2_e)$.

GEBVs were subsequently used to calculate a superior progeny value (Zhong and Jannink, 2007) from all desirable 1,100 crosses using the R package PopVar (Mohammadi et al, 2015). For each genetic architecture, three sets of 300 crosses (designated GA, GB, and GC) with the greatest superior progeny values were used to initiate the variety development. Subsequent development of ten varieties from each set followed the same process as that of the second and third experiments.

Simulation experiment 5: Relative predictions +MAS+genotypic selection.

The same three sets (GA, GB and GC) that were used to initiate the variety development project in experiment 4 were used. As with experiments 2, 3 and 4 MAS for non-yield traits was conducted in the F_2 generation. Instead of using phenotypic selection for yield in the PYT stage,

GEBV's for yield were used to select 10% of the 15,375 lines evaluated in the PYT. As with experiment 4 data from a previous cycle of variety development were used as a training set to calculate GEBVs for the lines evaluated in the PYT. The model used to calculate GEBV's was based on the genotypes and phenotypes of selected lines evaluated in the AYT trials of experiment 3 as training sets. In other words, training sets consisted of genotypic assays at 1,000 loci in combination with the simulated 1,540 lines evaluated in AYT-1 and 300 lines evaluated in AYT-2 responsible for yield values for each of the nine genetic architectures in each of the three sets of families (YA, YB and YC) to obtain GEBVs (Endelman, 2011) for 15,375 lines from each set (GA, GB and GC) evaluated in the PYT. Subsequent stages of variety development was based on phenotypic selection for yield, i.e., the same as with all previous experiments.

Evaluation metrics

Analyses of variance. There are a number of factors that could influence the outcomes from each experiment. These include the starting sets of 300 crosses, stage of selection, genetic architecture and heritability of the traits. These were allocated to the simulations in a split plot model in which sets of crosses are considered as whole plots and the remaining factors represent a full factorial set of 27 treatments. Subsequent to analyses of variance, specific contrasts of interest for any interactions among factors with demonstrated significant variability were compared using Tukey's *hsd*. If the differences between least squared mean (LSM) values exceeded a threshold associated with a type I error having probability < *0.01*, it was considered significant.

Effectiveness metrics. For purposes of evaluating selection decision accuracies, sensitivities and specificities, the genotypic values for yield represents the genetic potential of each line. Thus, we evaluated genotypic values comprised of only the genotypic component of the model, i.e., $\boldsymbol{y_i} = \boldsymbol{\mu} + \boldsymbol{G_i}$ for analyses of variance. Estimates of accuracy of GEBVs in experiment 5 utilized five-

fold cross validation. Ten replicates of the five-fold cross validation were used to estimate the prediction accuracy as Pearson's correlation between simulated phenotypes and GEBV values.

Decisions to select and discard lines represents a binary decision and since the data were simulated, the true genotypic values of all individuals and lines are known. Thus, decision accuracy, composed of sensitivity and specificity represent normalized metrics that can be determined using confusion tables (Table 6) to compare effectiveness of selection strategies.

Table 6: Example of a confusion table to determine decision accuracy, sensitivity and specificity.

N = Σ *lines*	**Yield or GEBV* (Discard)**	**Yield or GEBV (Select)**
Genotypic Yield (Discard)	TD	FS
Genotypic Yield (Select)	FD	TS

***** For experiments 1, 2, 3, and 4 only the simulated phenotypic yield was used to compare with genotypic values for yield; for experiment 5, GEBV was also used for analysis
TD = true discard; FS = false select; FD = false discard; TS = true select

Decision accuracy is the proportion of true selected (TS) and true discarded (TD) lines relatively to the total number of decisions:

$$Decision\ \ Accuracy = \frac{\boldsymbol{\Sigma\, TS + \Sigma\, TD}}{\boldsymbol{\Sigma\, total}}$$

Sensitivity is the proportion of correctly selected lines relative to all lines that should have been selected.

$$Sensitivity = \frac{\boldsymbol{\Sigma\, TS}}{\boldsymbol{\Sigma\, TS + \Sigma\, FD}}$$

Specificity is the proportion of correctly discarded lines relative to all lines that should have been discarded.

$$Specificity = \frac{\boldsymbol{\Sigma\, TD}}{\boldsymbol{\Sigma\, TD + \Sigma\, FS}}$$

Selection ability: Of interest is whether the ten best lines that enter into PYT stage survived to the EYT stage. This was evaluated for all five experiments with all nine genetic architecture and heritability combinations.

Cost effectiveness: initially, costs associated with experiments 2, 3, 4 and 5 were treated as additional costs relative to experiment 1. By establishing variety development costs using only phenotypic selection, i.e., experiment 1, it was possible to contrast cost effectiveness among the experiments and to generate and test hypotheses about trade-offs between costs and decision accuracies. Cost effectiveness among selection strategies was measured by the total budget (in $US) needed to implement each experiment by counting the number of plot units required. Tasks associated with each stage were assigned respective direct costs of $20 per plot and indirect costs of 50% per plot. Therefore, total unit-plots were estimated as the activity unit cost divided by $30. The plot costs for experiment 1 contains 23k plots with a cost of $690k, experiments 2,3, and 4 had 37k plot units and cost $1.1M, and experiment 5 required 47k plot units at a cost of $1.4M (Table 7).

Table 7: Resources allocation in terms of plots and US dollars for five experiments.

Generation	Task	Plots	Unit ($)	Experiment 1		Experiments 2-3-4		Experiment 5	
				$	Plots	$	Plots	$	Plots
F_1	selfing	300	7.5	2,250	75	2,250	75	2,250	75
F_2	selfing	300	30	9,000	300	9,000	300	9,000	300
F_2	SNPs	900,000	0.5			450,000	15,000	450,000	15,000
F_3	selfing	300	30	9,000	300	9,000	300	9,000	300
F_4	selfing	300	30	9,000	300	9,000	300	9,000	300
PYT	genotyping	15,375	20					307,500	10,250
PYT	field trial	15,375	7.5	115,313	3,844	115,313	3,844	115,313	3,844
AYT-1	field trial	7,700	30	231,000	7,700	231,000	7,700	231,000	7,700
AYT-2	field trial	7,200	30	216,000	7,200	216,000	7,200	216,000	7,200
AYT-2	phenotyping	300	87	26,100	870				
EYT	field trial	2,400	30	72,000	2,400	72,000	2,400	72,000	2,400
Total budget				**689,663**	**22,989**	**1,113,563**	**37,119**	**1,421,063**	**47,369**

The number of families, plants and lines are the same for all five experiments. F_1 costs were estimated at \$7.5 or 0.25 per unit-plot. F_2, F_3, F_4, were given the same cost at \$30 or 1 unit-plot for self-pollination. The cost of selecting individuals carrying favorable alleles among F_2 individuals in experiments 2, 3, 4, and 5 was \$0.35/sample of DNA extraction and \$0.05/assay in a total of \$0.5/plant or 0.017 unit-plot. For experiments 4 and 5, GEBVs were determined to select crosses to make, so genotyping costs had to be applied to the lines used as founder crosses. For experiment 5, each one of the 15,375 PYTs was fingerprinted for a cost of \$20/PYT or 0.67 unit-plot (http://soybase.org/) assuming that Illumina Soy6k SNP Infinium chip cost \$20-25/sample. Phenotypic values for the training set are typically amortized from AYTs and EYT stages of previous variety development projects. For these simulations the phenotypic and genotypic values from experiment 3 were used, so the costs associated with creating the training sets are directly transferable. 1,848 lines selected from the PYTs had seeds multiplied, and after harvesting 308 random lines were discarded resulting in 1,540 lines for AYT-1. Each grown PYT has a cost of \$7.5 and 0.25 unit-plot. AYT-1, AYT-2, and EYT field trial plot cost \$30. Phenotypic screening at AYT-2 stage was \$50/line of one SCN race and \$37/line for PRR and 2.9 unit-plots/line.

Costs associated crossing (Table 4) were not shown since they were the same for all experiments. Relative efficiencies were determined among the combinations of experiments. Explicitly break-even for cost and number of plots between experiments was determined by genotyping costs for the lines used as founder crosses, estimating the F_2 population size, SCN and PRR greenhouse phenotyping in AYT-2 and genotyping of PYTs.

RESULTS

Genotypic data

After elimination of non informative SNP markers, each of the112 potential founder lines were genotypically characterized with 34,708 SNPs (Table 8).

Table 8: Example of genotypic data from 112 founder lines.

Founders / SNPs	PI615555	PI548655	PI614088	PI620883		PI634335
1	AA	GG	AA	AA	...	GG
2	CC	TT	CC	CC		TT
3	AA	GG	AA	AA		GG
4	CC	TT	CC	CC	...	TT
5	AA	GG	AA	AA		GG
...			...			
34,708	GG	AA	GG	GG	...	AA

Cluster analyses based on genetic similarities (Nei and Li, 1979) estimated with the 34,708 SNP markers among all pairs of 112 lines revealed five groups (Figure 6).

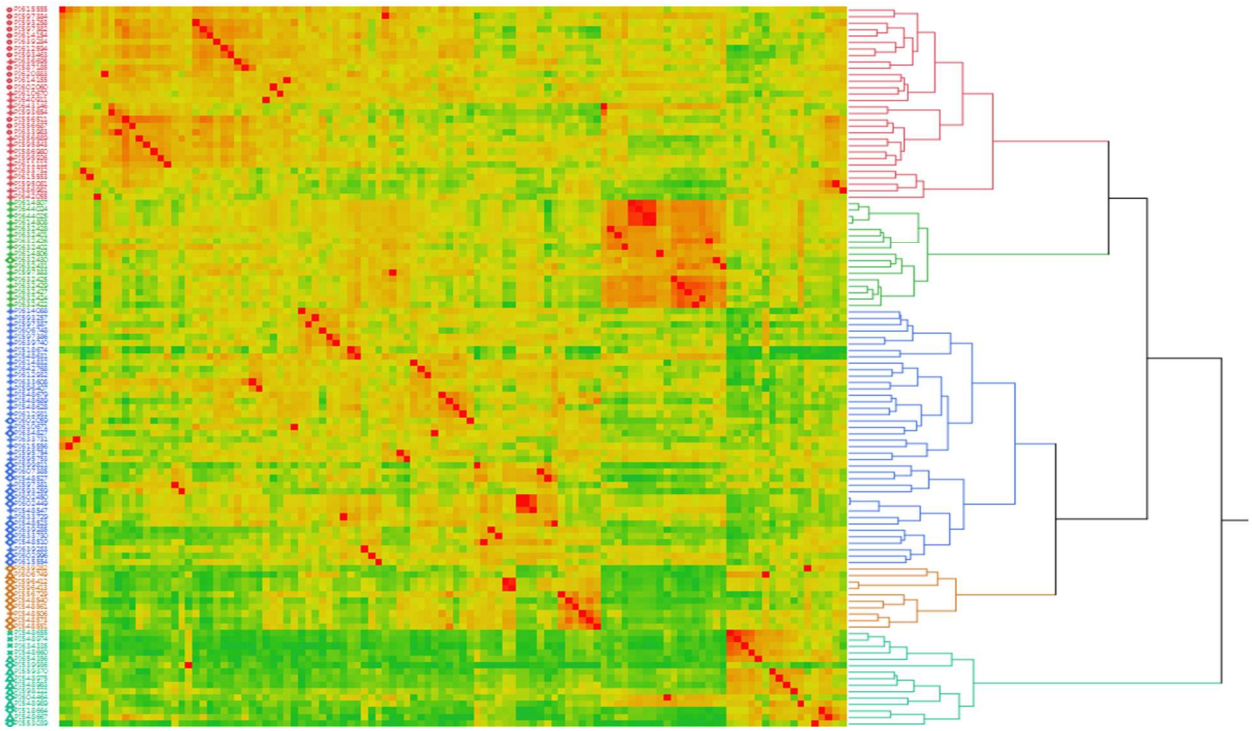

Figure 6: Hierarchical cluster showing genetic similarity among soybean lines RM I – VII.

RM values of members in each cluster indicates that the first, second and third groups consist of accessions adapted to MZs II, III and IV. The fourth cluster has most members adapted to the early MZs and the fifth cluster consists of members adapted to late MZs (Table 9).

Table 9: Membership in five clusters of 112 soybean accessions and their respective RMs.

Cluster	RM							
	I	II	III	IV	V	VI	VII	Total
1		5	12	13				30
2		1	9	7				17
3		13	19	8				40
4	2	6	1	1				10
5				5	5	3	2	15

Simulated phenotypes of the founders and their progeny

Simulated yields of potential founders ranged from 48.9 – 74.7 bu/ac, RM values ranged from I – VII, while lodging, SCN, and PRR ordinal scores ranged from 1 to 9 (Table 10). The distributions for each of the simulated traits at each of the stages of varietal development indicate that implementation of genetic architectures for the simulation models similar to those observed in actual soybean variety development projects (Table 5).

Table 10: Representation of founder phenotype file for yield, RM, lodging, SCN and PRR traits.

Founder	Yield	RM	Lodging	SCN	PRR
PI615555	59.43	III	6	1	9
PI548655	63.44	V	7	9	1
PI614088	63.64	II	4	1	3
PI620883	69.44	IV	5	1	7
PI634335	62.82	IV	7	9	1
. .	.	. .	. .	. .	. .
PI539936	63.11	IV	1	1	3

Simulation experiment 1: phenotypic selection

A comparison of founding crosses selected for experiment 1 indicated that 62% were in common among the three random sets. Analysis of variance of the simulated genotypic values for yield (Appendix A3) revealed that outcomes were due to interactions between genetic architectures and heritabilities and their interactions with either founder sets or stages of development. In other words, selection based on phenotypic yield beginning in the PYT stage and subsequent evaluation stages (AYT-1, AYT-2 and EYT) were dependent upon which random set of 300 crosses from among the 1,100 were used and the simulated heritabilities nested within genetic architectures for yield (Appendices A4-A6). Although the specific sets of lines and yield values were not consistent across all subsequent stages of evaluation, the distributions of the selected lines' phenotypes were similar (Figure 7).

Analysis of variance of EYT lines revealed the simulated genotypic values for yield (Table 11) were affected by interactions with heritability, genetic architecture and founder sets.

Table 11: Analysis of variance for simulated genotypic values of yield due to founder sets, genetic architectures and heritability for lines in the final stage (EYT) of variety development in experiment 1 where phenotypic selection was conducted on all traits.

Source	**Sum of Squares**	**Mean Square**	**F Ratio**	**Prob > F**
Model	58314.8	2242.8	4984.6	<.0001*
Error	7277.1	0.4		
C. Total	65592.0			
Source		**S.Squares**	**F Ratio**	**Prob > F**
FounderSet		132.8	147.6	<.0001*
GeneticArchitecture		20365.5	22630.4	<.0001*
FounderSet*GeneticArchitecture		73.3	40.7	<.0001*
Heritability		34425.9	38254.5	<.0001*
FounderSet*Heritability		223.6	124.2	<.0001*
GeneticArchitecture*Heritability		2923.7	1624.4	<.0001*
FounderSet*GeneticArchitecture*Heritability		169.7	47.1	<.0001*

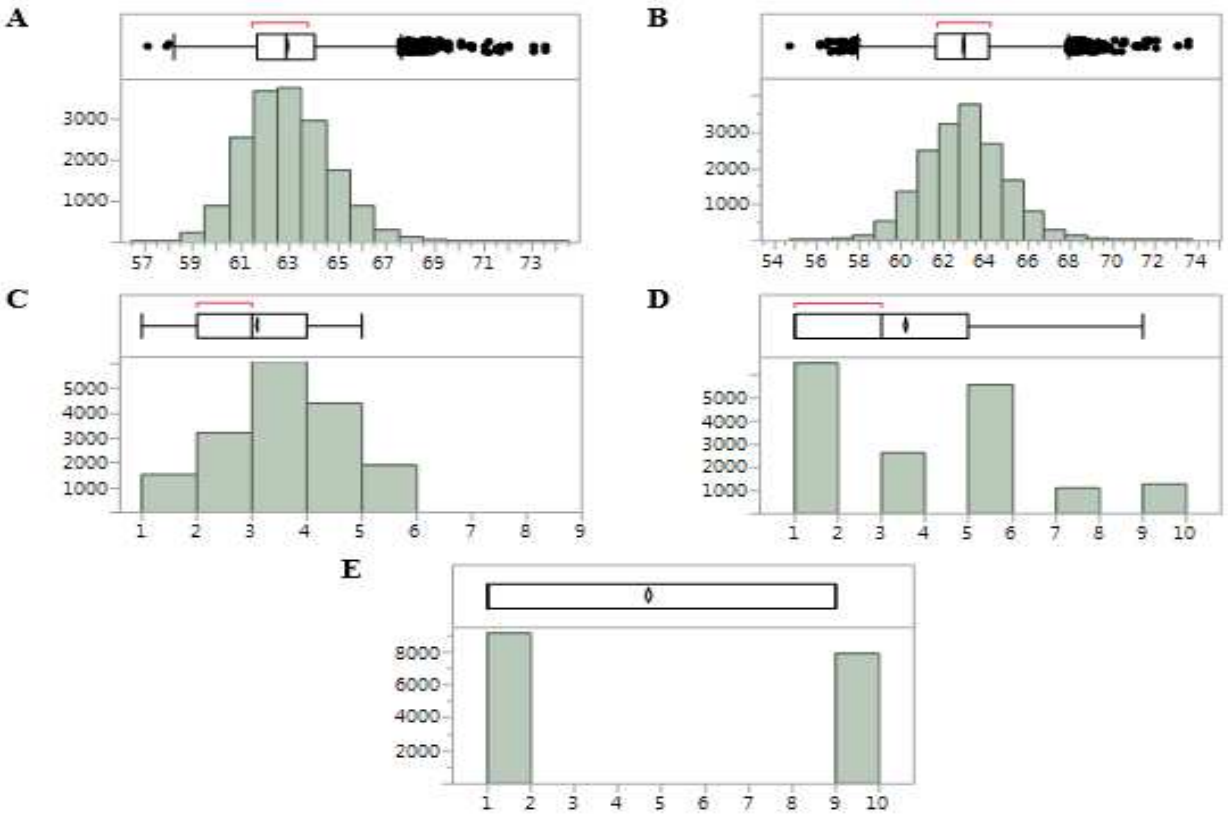

Figure 7: Distributions of simulated A) genotypic values of yield, B) phenotypic values of yield, C) lodging, D) PRR, and E) SCN of lines evaluated from experiment 1. Distributions are composed of data aggregated from lines across all stages of development and all combinations of simulated genetic architectures and heritabilities.

Phenotypic data for AYT-1, AYT-2 and EYT testing stages are summarized as the average of selected individuals for heritabilities nested within genetic architectures (Appendices A4-A6). AYT-1 $F_{4:5}$ lines (75 by RM group) were selected for each heritability within genetic architecture and evaluated in AYT-2. Average genotypic values for yield ranged from 61.6-67.2 bu/ac, phenotypic yield 62.4-67.2 bu/ac, lodging 2.6-3.6, SCN 2.6-7.4, and PRR 3.0-4.8. Subsequently, for each combination of heritability within genetic architecture, 40 AYT-2 $F_{4:7}$ lines (10 by RM group) were selected for evaluation in EYT. Average simulated genotypic values of yield ranged from 61.4-69.1 bu/ac, average phenotypic yield 61.6-69.1 bu/ac, lodging 2.2-4.0, SCN 1.0, and PRR 1.0-3.0. Last, 10 high yield EYT $F_{4:7}$ lines had the attributes to be

released as new variety. Simulated genotypic values for yield ranged from 61.7-70.3 bu/ac, phenotypic yield ranged from 62.5-70.3 bu/ac, lodging 1.0-5.0, SCN 1.0, and PRR 1.0-3.0.

There were significant differences between averages of simulated genotypic values and phenotypic yield values on an entry mean basis, depending on combinations of heritability and genetic architecture, stages of development and RM groups (Appendices A7-A9). There were significant differences (Tukey α=0.01) observed for simulated phenotypic yield values, on an entry mean basis, among stages of development and RM groups, where the averaged phenotypic yields of lines in EYTs were greater than the averaged phenotypic yields of lines in the AYT's (Table 12). Recall yield and RM were not simulated using a model with correlated genotypic values between these two traits. Significant differences were observed for SCN and PRR between AYT-1 and the other two stages of development. There were no significant differences observed for lodging (Tukey α=0.01), except for values in RM 3.0-3.6.

Table 12: Genotypic values for yield and least square means for simulated yield, lodging, SCN and PRR in experiment 1. Values are listed by RM group and varietal development stages.

Stage	RM group	Genotypic yield mean	Phenotypic yield mean	Lodging mean	SCN mean	PRR mean
AYT-1	1.8-2.4	63.5 ^{A}	64.3 ^{A}	3.0 ^{A}	5.0 ^{A}	3.7 ^{A}
AYT-2	1.8-2.4	64.0 ^{B}	64.2 ^{A}	3.1 ^{A}	1.0 ^{B}	1.7 ^{B}
EYT	1.8-2.4	64.7 ^{C}	65.2 ^{B}	3.3 ^{A}	1.0 ^{B}	1.6 ^{B}
AYT-1	2.4-3.0	63.9 ^{A}	64.7 ^{A}	3.1 ^{A}	4.8 ^{A}	3.9 ^{A}
AYT-2	2.4-3.0	64.3 ^{B}	64.6 ^{A}	3.2 ^{A}	1.0 ^{B}	1.7 ^{B}
EYT	2.4-3.0	65.0 ^{C}	65.4 ^{B}	3.3 ^{A}	1.0 ^{B}	1.5 ^{B}
AYT-1	3.0-3.6	64.1 ^{A}	64.9 ^{A}	3.1 ^{A}	4.9 ^{A}	3.8 ^{A}
AYT-2	3.0-3.6	64.5 ^{B}	64.7 ^{A}	3.3 ^{B}	1.0 ^{B}	1.7 ^{B}
EYT	3.0-3.6	65.2 ^{B}	65.6 ^{B}	3.5 ^{B}	1.0 ^{B}	1.5 ^{B}
AYT-1	3.6-4.2	64.1 ^{A}	64.8 ^{A}	3.2 ^{A}	4.6 ^{A}	3.7 ^{A}
AYT-2	3.6-4.2	64.5 ^{B}	64.8 ^{A}	3.2 ^{A}	1.0 ^{B}	1.6 ^{B}
EYT	3.6-4.2	65.3 ^{B}	65.8 ^{B}	3.2 ^{A}	1.0 ^{B}	1.6 ^{B}

Levels not connected by the same letter are significantly different (Tukey α=0.01)

While the sets of lines selected from the PYTs were distinct among the combinations of genetic architecture and heritability, there was little evidence that decision accuracy and

Table 13: Average decision accuracy, sensitivity, and specificity from three founder sets used in experiment 1 for nine combinations of genetic architecture and heritability for each of four stages in soybean variety development.

Stage	Yield genetic architecture	H^2	Accuracy	(SE)	Sensitivity	(SE)	Specificity	(SE)
PYT	3-categories QTL	0.2	0.88	(±0.01)	0.39	(±0.01)	0.93	(±0.01)
		0.6	0.92	(±0.01)	0.62	(±0.01)	0.96	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Uniform	0.2	0.88	(±0.01)	0.38	(±0.01)	0.93	(±0.01)
		0.6	0.93	(±0.01)	0.66	(±0.02)	0.96	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Negative Exponential	0.2	0.88	(±0.01)	0.40	(±0.01)	0.93	(±0.01)
		0.6	0.93	(±0.01)	0.64	(±0.01)	0.96	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
AYT-1	3-categories QTL	0.2	0.79	(±0.01)	0.47	(±0.01)	0.87	(±0.01)
		0.6	0.85	(±0.01)	0.62	(±0.02)	0.91	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Uniform	0.2	0.82	(±0.01)	0.54	(±0.01)	0.89	(±0.01)
		0.6	0.88	(±0.01)	0.69	(±0.02)	0.92	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Negative Exponential	0.2	0.81	(±0.01)	0.51	(±0.01)	0.88	(±0.01)
		0.6	0.86	(±0.01)	0.64	(±0.01)	0.91	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
AYT-2	3-categories QTL	0.2	0.92	(±0.01)	0.71	(±0.02)	0.96	(±0.01)
		0.6	0.95	(±0.01)	0.80	(±0.02)	0.97	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Uniform	0.2	0.92	(±0.01)	0.68	(±0.01)	0.95	(±0.01)
		0.6	0.97	(±0.01)	0.90	(±0.02)	0.98	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Negative Exponential	0.2	0.95	(±0.02)	0.83	(±0.06)	0.97	(±0.01)
		0.6	0.94	(±0.01)	0.79	(±0.04)	0.97	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
EYT	3-categories QTL	0.2	0.77	(±0.02)	0.53	(±0.04)	0.84	(±0.02)
		0.6	0.78	(±0.02)	0.57	(±0.04)	0.86	(±0.02)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Uniform	0.2	0.73	(±0.02)	0.47	(±0.04)	0.82	(±0.02)
		0.6	0.80	(±0.03)	0.60	(±0.06)	0.87	(±0.02)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Negative Exponential	0.2	0.73	(±0.04)	0.47	(±0.07)	0.82	(±0.03)
		0.6	0.80	(±0.03)	0.60	(±0.06)	0.87	(±0.02)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)

specificity of selections for yield were affected by the founder sets (Table 13). Results for each set are shown in Appendices A10-A12. As a reference, if all of the phenotypic variance can be

explained by the genotypic variance, decision accuracy, sensitivity and specificity are equal to 1.00. Overall, high values were observed for decision accuracy due primarily to the high specificity, i.e., ability to discard lines that should be discarded. On the other hand, decision sensitivity, i.e., ability to retain lines that should be retained was not consistent among combinations of genetic architecture and heritability. Higher values observed for AYT-2, can be explained by the selection applied at this stage for SCN and PRR resistant individuals, resulting in discarding lines with high yield.

Simulation experiment 2: MAS for SCN and PRR traits + phenotypic selection for yield

As with experiment 1, analysis of variance of the simulated yield values (Appendix A13) revealed that yields of the set of lines selected at each stage of development for each simulated level of heritability within the genetic architectures were not consistent across the three founder sets. Also, analysis of variance for EYT lines revealed the simulated genotypic values for yield was affected by interactions with heritability, genetic architecture and founder sets (Table 14).

Table 14: Analysis of variance for simulated genotypic values for yield in the final stage (EYT) of variety development due to founder sets, genetic architectures and heritability of EYT lines for experiment 2 where MAS for SCN and PRR traits was conducted prior to the PYT and phenotypic selection for yield was conducted through all stages of development.

Source	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	139577.8	5368.3	7832.1	<.0001*
Error	22189.2	0.6		
C. Total	161767.1			
Source		**S.Squares**	**F Ratio**	**Prob > F**
FounderSet		2818.4	2055.9	<.0001*
GeneticArchitecture		42519.8	31017.0	<.0001*
FounderSet*GeneticArchitecture		562.2	205.0	<.0001*
Heritability		54876.4	40030.9	<.0001*
FounderSet*Heritability		602.1	219.6	<.0001*
GeneticArchitecture*Heritability		1797.1	655.4	<.0001*
FounderSet*GeneticArchitecture*Heritability		554.9	101.1	<.0001*

Although the specific sets of lines and yield values were not consistent across all subsequent stages of evaluation, the distributions of the selected lines' phenotypes were similar (Figure 8). Phenotypic data for AYT-1, AYT-2 and EYT testing stages are summarized as the average of selected individuals for heritabilities nested within genetic architectures (Appendices A14-A16). AYT-1 $F_{4:5}$ lines (75 by RM group) were selected for each heritability within genetic architecture and evaluated in AYT-2. Average genotypic values for yield ranged from 61.6-67.2 bu/ac, phenotypic yield 62.4-67.2 bu/ac, lodging 2.5-3.6, SCN 1.0, and PRR 1.0-1.3.

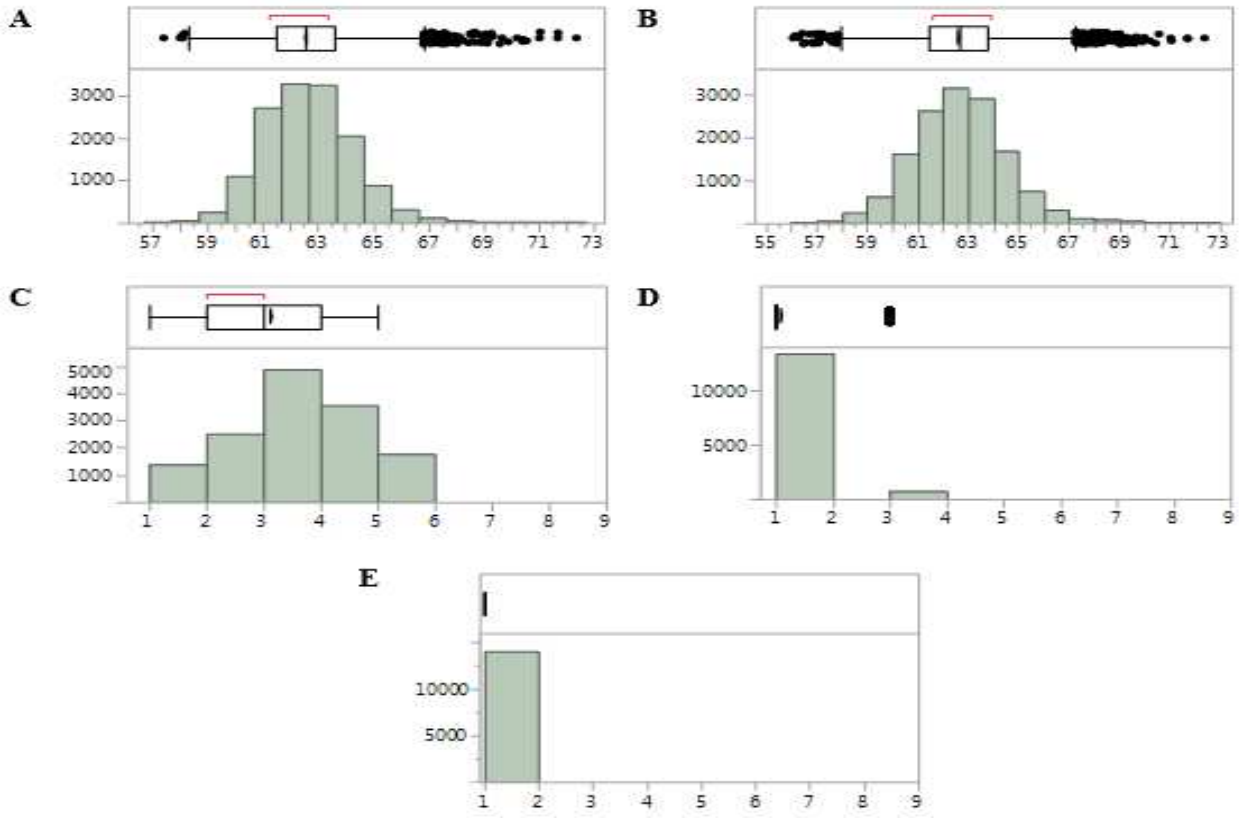

Figure 8: Distributions of simulated A) genotypic values of yield, B) phenotypic values of yield, C) lodging, D) PRR, and E) SCN of lines evaluated from experiment 2. Distributions are composed of data aggregated from lines selected for all combinations of simulated genetic architectures and heritabilities.

Subsequently, for each heritability genetic architecture combination, 40 AYT-2 $F_{4:7}$ lines (10 by RM group) were selected for evaluation in EYT. Average simulated genotypic values of

yield ranged from 62.0-70.2 bu/ac, average phenotypic yield 62.8-70.2 bu/ac, lodging 2.2-4.3, SCN 1.0, and PRR 1.0-1.6. Last, 10 high yield EYT $F_{4:7}$ lines had the attributes to be released as new variety. Simulated genotypic values for yield ranged from 62.1-71.7 bu/ac, simulated phenotypic values for yield ranged from 62.9-71.7 bu/ac, lodging 1.0-5.0, SCN 1.0, and PRR 1.0-3.0.

There were significant differences between averages of simulated genotypic values and averages of phenotypic yield values on an entry mean basis, depending on combinations of heritabilities and genetic architectures, stages of development and RM groups (Appendices A17-A19). Overall, genotypic and phenotypic yield values in AYT-2 and EYT were consistently higher than yield in AYT-1 (Table 15). Significant differences were not observed for lodging, SCN, and PRR among stages of development and RM groups.

Table 15: Genotypic values for yield and least square means for simulated yield, lodging, SCN and PRR in experiment 2. Values are listed by RM group and varietal development stages.

Stage	RM group	Genotypic yield mean	Phenotypic yield mean	Lodging mean	SCN mean	PRR mean
AYT-1	1.8-2.4	63.5 ^{A}	64.3 ^{A}	3.1 ^{A}	1.0 ^{A}	1.1 ^{A}
AYT-2	1.8-2.4	64.8 ^{B}	65.3 ^{B}	3.1 ^{A}	1.0 ^{A}	1.1 ^{A}
EYT	1.8-2.4	65.4 ^{B}	66.0 ^{B}	3.1 ^{A}	1.0 ^{A}	1.1 ^{A}
AYT-1	2.4-3.0	63.8 ^{A}	64.5 ^{A}	3.1 ^{A}	1.0 ^{A}	1.2 ^{A}
AYT-2	2.4-3.0	65.0 ^{B}	65.5 ^{B}	3.2 ^{A}	1.0 ^{A}	1.1 ^{A}
EYT	2.4-3.0	65.5 ^{B}	66.0 ^{B}	3.2 ^{A}	1.0 ^{A}	1.1 ^{A}
AYT-1	3.0-3.6	64.0 ^{A}	64.7 ^{A}	3.1 ^{A}	1.0 ^{A}	1.2 ^{A}
AYT-2	3.0-3.6	65.3 ^{B}	65.7 ^{B}	3.2 ^{A}	1.0 ^{A}	1.1 ^{A}
EYT	3.0-3.6	66.0 ^{B}	66.6 ^{C}	3.4 ^{A}	1.0 ^{A}	1.1 ^{A}
AYT-1	3.6-4.2	64.0 ^{A}	64.7 ^{A}	3.1 ^{A}	1.0 ^{A}	1.1 ^{A}
AYT-2	3.6-4.2	65.4 ^{B}	65.7 ^{B}	3.2 ^{A}	1.0 ^{A}	1.1 ^{A}
EYT	3.6-4.2	66.3 ^{C}	66.6 ^{C}	3.2 ^{A}	1.0 ^{A}	1.1 ^{A}

Levels not connected by the same letter are significantly different (Tukey α=0.01)

While the sets of lines selected from the PYTs were distinct among the genetic architecture – heritability combinations, there was little evidence that these affected decision accuracy and specificity of selections for yield (Table 16).

Table 16: Average decision accuracy, sensitivity, and specificity from three founder sets used in experiment 2 for nine combinations of genetic architecture and heritability for each of four stages in soybean variety development.

Stage	Yield genetic architecture	H^2	Accuracy	(SE)	Sensitivity	(SE)	Specificity	(SE)
PYT	3-categories QTL	0.2	0.87	(±0.01)	0.39	(±0.01)	0.93	(±0.01)
		0.6	0.92	(±0.01)	0.63	(±0.01)	0.96	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Uniform	0.2	0.87	(±0.01)	0.39	(±0.01)	0.93	(±0.01)
		0.6	0.93	(±0.01)	0.67	(±0.01)	0.96	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Negative Exponential	0.2	0.87	(±0.01)	0.40	(±0.01)	0.93	(±0.01)
		0.6	0.92	(±0.01)	0.63	(±0.01)	0.96	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
AYT-1	3-categories QTL	0.2	0.80	(±0.01)	0.48	(±0.01)	0.87	(±0.01)
		0.6	0.84	(±0.01)	0.60	(±0.01)	0.90	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Uniform	0.2	0.81	(±0.01)	0.52	(±0.01)	0.88	(±0.01)
		0.6	0.87	(±0.01)	0.68	(±0.03)	0.92	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Negative Exponential	0.2	0.83	(±0.01)	0.56	(±0.01)	0.89	(±0.01)
		0.6	0.87	(±0.01)	0.66	(±0.02)	0.92	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
AYT-2	3-categories QTL	0.2	0.87	(±0.01)	0.52	(±0.03)	0.93	(±0.01)
		0.6	0.90	(±0.01)	0.63	(±0.02)	0.94	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Uniform	0.2	0.87	(±0.00)	0.50	(±0.00)	0.92	(±0.01)
		0.6	0.91	(±0.01)	0.67	(±0.02)	0.95	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Negative Exponential	0.2	0.85	(±0.01)	0.43	(±0.04)	0.91	(±0.01)
		0.6	0.92	(±0.01)	0.72	(±0.01)	0.96	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
EYT	3-categories QTL	0.2	0.73	(±0.02)	0.47	(±0.04)	0.82	(±0.02)
		0.6	0.78	(±0.02)	0.57	(±0.04)	0.86	(±0.02)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Uniform	0.2	0.80	(±0.01)	0.60	(±0.00)	0.87	(±0.00)
		0.6	0.80	(±0.03)	0.60	(±0.06)	0.87	(±0.02)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Negative Exponential	0.2	0.72	(±0.02)	0.43	(±0.04)	0.81	(±0.02)
		0.6	0.80	(±0.03)	0.60	(±0.06)	0.87	(±0.02)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)

Results for each set are shown in Appendices A20-A22. Similar to experiment 1, if all of the phenotypic variance can be explained by the genotypic variance, decision accuracy,

sensitivity and specificity were equal to 1.00. Overall, high values were observed for decision accuracy due primarily to the ability to discard lines that should be discarded. Decision sensitivity was more affected by the combination of heritability and genetic architecture. Unlike experiment 1, where sensitivity values were higher in AYT-2 due to SCN and PRR selection, the differences observed for experiment 2 were more consistent among stages, once selection for SCN and PRR was conducted in F_2.

Simulation experiment 3: Selections of crosses based on performance of 112 potential founders + MAS for SCN and PRR traits in early generations + phenotypic selection for yield through all stages of development.

A high proportion of selected founding crosses used to initiate variety development in experiment 3 were with in common (Table 17). Also, a comparison of the initial crosses used to initiate variety development in experiment 3 with initial crosses used to initiate variety development in experiments 1 and 2 indicates that 56% were in common.

Table 17: Number of selected crosses that were consistent among three sets of 300 selected crosses. Selected sets of crosses were based on predicted GEBVs of 200 simulated recombinant inbred lines derived from 1,100 possible crosses that could be used to initiate a variety selection program.

Genetic architecture	Set	Number of crosses that were consistent with the first set	%
3-categories QTL	1	300	100
	2	299	99.7
	3	299	99.7
Uniform	1	300	100
	2	295	98.3
	3	275	90.7
Negative Exponential	1	300	100
	2	294	98.0
	3	296	98.7

Analysis of variance of the simulated genotypic values for yield revealed that yield was affected by interactions involving combinations of genetic architecture and heritability with the founding sets of lines as well with stages of development (Appendix A23). These interactions

make sense because the set of lines selected from the PYT for each simulated level of heritability were not consistent across the genetic architectures and stages of selection. As a consequence, simulated yield phenotypes for lines selected for subsequent evaluation stages (AYT-1, AYT-2 and EYT) were dependent upon the initial set of lines selected from the PYT (Appendices A24-A26).

Analysis of variance revealed the simulated genotypic values for yield in the EYT stage (Table 18) was affected by interactions with heritability, genetic architecture and founder sets of lines as well with stages of development (Appendix A23).

Table 18: Analysis of variance for simulated genotypic values for yield due to founder sets, genetic architectures and heritability of lines in experiment 3, in which founder crosses were selected based on training sets consisting of 1,000 SNP assays and simulated yields in 112 founders plus MAS for SCN and PRR in early generations and phenotypic selection for yield in subsequent stages of field trials.

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	26	86884.1	3341.7	6927.4	<.0001*
Error	16173	7801.6	0.4		
C. Total	16199	94685.7			

Source	DF	S.Squares	F Ratio	Prob > F
FounderSet	2	110.1	114.1	<.0001*
GeneticArchitecture	2	34406.5	35662.7	<.0001*
FounderSet*GeneticArchitecture	4	144.7	74.9	<.0001*
Heritability	2	50594.5	52441.7	<.0001*
FounderSet*Heritability	4	36.1	18.7	<.0001*
GeneticArchitecture*Heritability	4	1404.3	727.8	<.0001*
FounderSet*GeneticArchitecture*Heritability	8	187.7	48.6	<.0001*

Although the specific lines were not consistent across all subsequent stages of evaluation, the distributions of the phenotypes (Figure 9) were very similar to those generated in experiments 1 and 2. Simulated genotypic values for yield of lines selected from AYT-1 ranged from 61.5-67.5 bu/ac, while simulated phenotypic yield on an entry mean basis ranged from 62.3-67.5 bu/ac, lodging 2.6-3.7, SCN 1.0, and PRR 1.0-1.3. Simulated genotypic values for yield for 40 lines selected from AYT-2 ranged from 62.0-70.0 bu/ac, averaged phenotypic yield

values ranged from 62.7-70.0 bu/ac, lodging 2.2-4.2, SCN 1.0, and PRR 1.0-1.8. Last, 10 high yield EYT $F_{4:7}$ lines had the attributes to be released as new variety. Simulated genotypic values for yield ranged from 62.0-71.3 bu/ac, averaged phenotypic yield ranged from 62.6-71.3 bu/ac, lodging 1.0-5.0, SCN 1.0, and PRR 1.0-2.3.

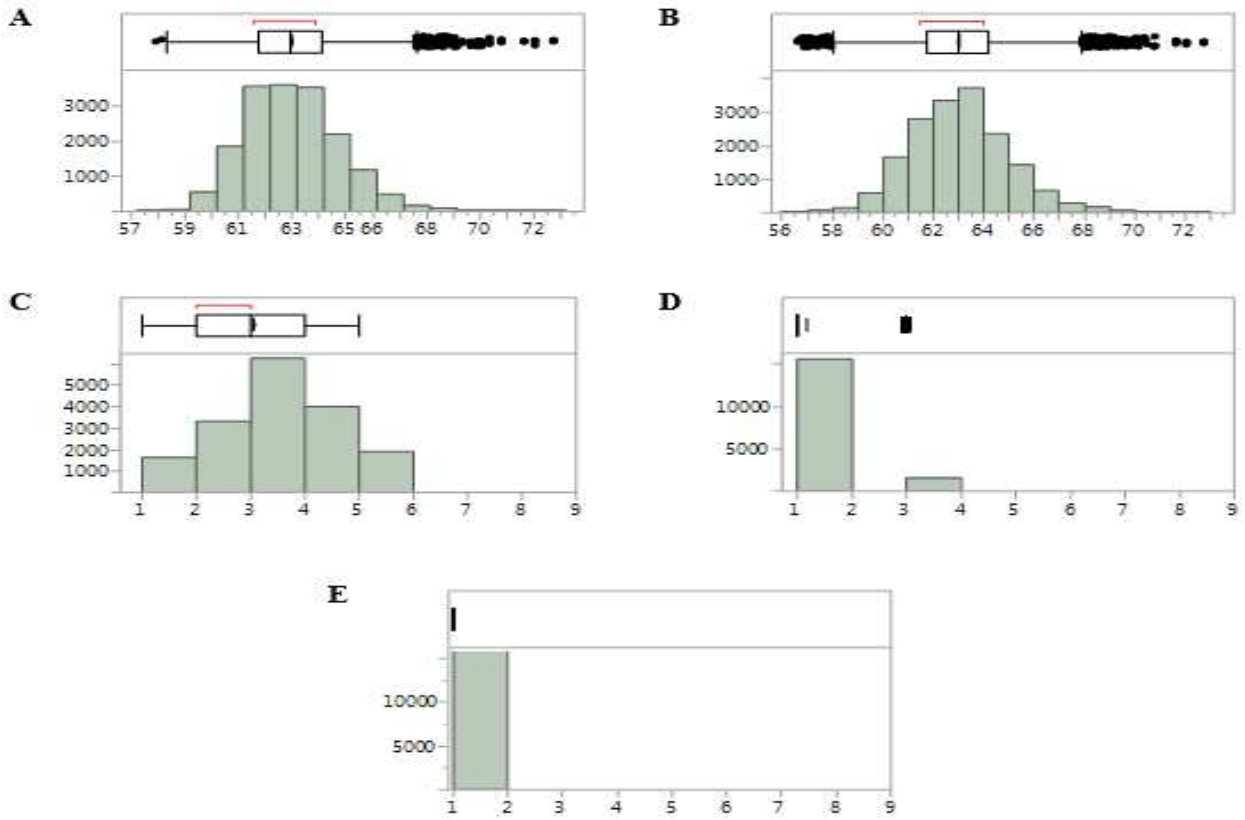

Figure 9: Distributions of simulated A) genotypic values of yield, B) phenotypic values of yield, C) lodging, D) PRR, and E) SCN of lines evaluated from experiment 3. Distributions are composed of data aggregated from lines selected for all combinations of simulated genetic architectures and heritabilities.

There were significant differences between averages of simulated genotypic values and simulated phenotypic yield on an entry mean basis among genetic architectures, stages of development and RM groups (Appendices A27-A29). Overall, genotypic and phenotypic yield values in AYT-2 and EYT were consistently higher than these values in AYT-1 (Table 19). As with experiment 2, significant differences were not observed for lodging, SCN, and PRR among stages of development and RM groups.

Table 19: Genotypic values for yield and least square means for simulated yield, lodging, SCN and PRR in experiment 3, listed by RM group and varietal development stages in which founder crosses were selected based on training sets consisting of 1,000 SNP assays and simulated yields in 112 founders plus MAS for SCN and PRR traits plus phenotypic selection for yield.

Stage	RM	Genotypic yield mean	Phenotypic yield mean	Lodging mean	SCN mean	PRR mean
AYT-1	1.8-2.4	63.7 ^{A}	64.4 ^{A}	3.1 ^{A}	1.0 ^{A}	1.2 ^{A}
AYT-2	1.8-2.4	65.0 ^{B}	65.4 ^{B}	3.1 ^{A}	1.0 ^{A}	1.1 ^{A}
EYT	1.8-2.4	65.3 ^{B}	65.9 ^{B}	2.9 ^{A}	1.0 ^{A}	1.1 ^{A}
AYT-1	2.4-3.0	64.0 ^{A}	64.7 ^{A}	3.1 ^{A}	1.0 ^{A}	1.2 ^{A}
AYT-2	2.4-3.0	65.2 ^{B}	65.7 ^{B}	3.1 ^{A}	1.0 ^{A}	1.2 ^{A}
EYT	2.4-3.0	65.7 ^{B}	66.2 ^{B}	3.2 ^{A}	1.0 ^{A}	1.1 ^{A}
AYT-1	3.0-3.6	64.3 ^{A}	65.0 ^{A}	3.1 ^{A}	1.0 ^{A}	1.2 ^{A}
AYT-2	3.0-3.6	65.5 ^{B}	66.0 ^{B}	3.0 ^{A}	1.0 ^{A}	1.2 ^{A}
EYT	3.0-3.6	66.2 ^{B}	66.5 ^{B}	3.0 ^{A}	1.0 ^{A}	1.2 ^{A}
AYT-1	3.6-4.2	64.2 ^{A}	64.9 ^{A}	3.2 ^{A}	1.0 ^{A}	1.2 ^{A}
AYT-2	3.6-4.2	65.5 ^{B}	65.9 ^{B}	3.2 ^{A}	1.0 ^{A}	1.2 ^{A}
EYT	3.6-4.2	66.2 ^{B}	66.7 ^{C}	3.1 ^{A}	1.0 ^{A}	1.2 ^{A}

Levels not connected by the same letter are significantly different (Tukey α=0.01)

Differences among decision accuracy, sensitivity and specificity of selected lines for yield are reported for all three sets of founder lines in which genetic architectures explain 20% and 60% of the phenotypic variance on an entry mean basis (Appendices A30-32). However, since there was little evidence for founder set effects on either averaged phenotypic yield values and genotypic values for yield, decision accuracies, sensitivities and specificities were determined for data pooled from all three founder sets (Table 20). As with experiments 1 and 2, decision accuracy, sensitivity and specificity were equal to 1.00 in cases where all of the simulated phenotypic variability is explained by simulated genotypic variability. Similar to experiments 1 and 2, high decision accuracy was realized at each selection stage which was due primarily to the high specificity values. Unlike experiment 1 and 2, AYT-2 stage values for sensitivity were similar to other stages. This was most likely because selection for SCN and PRR had been carried out using MAS prior to the PYT stage.

Table 20: Average decision accuracy, sensitivity, and specificity metrics from three sets of founder crosses for each heritability genetic architecture combination and stage of development in soybean variety development. Experiment 3 consists of founder crosses that were selected based on training sets consisting of 1,000 SNP assays and simulated yields in 112 founders plus MAS for SCN and PRR traits plus phenotypic selection for yield.

Stage	Yield genetic architecture	H^2	Accuracy	(SE)	Sensitivity	(SE)	Specificity	(SE)
PYT	3-categories QTL	0.2	0.88	(±0.01)	0.38	(±0.01)	0.93	(±0.01)
		0.6	0.92	(±0.01)	0.62	(±0.01)	0.96	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Uniform	0.2	0.88	(±0.01)	0.39	(±0.01)	0.93	(±0.01)
		0.6	0.93	(±0.01)	0.66	(±0.02)	0.96	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Negative Exponential	0.2	0.88	(±0.01)	0.38	(±0.01)	0.93	(±0.01)
		0.6	0.92	(±0.01)	0.61	(±0.01)	0.96	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
AYT-1	3-categories QTL	0.2	0.79	(±0.01)	0.46	(±0.02)	0.87	(±0.01)
		0.6	0.85	(±0.01)	0.61	(±0.01)	0.91	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Uniform	0.2	0.82	(±0.01)	0.54	(±0.02)	0.89	(±0.01)
		0.6	0.88	(±0.01)	0.68	(±0.02)	0.92	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Negative Exponential	0.2	0.82	(±0.01)	0.55	(±0.01)	0.89	(±0.01)
		0.6	0.86	(±0.01)	0.64	(±0.02)	0.91	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
AYT-2	3-categories QTL	0.2	0.85	(±0.01)	0.44	(±0.04)	0.91	(±0.01)
		0.6	0.92	(±0.01)	0.71	(±0.04)	0.96	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Uniform	0.2	0.86	(±0.02)	0.46	(±0.06)	0.92	(±0.01)
		0.6	0.90	(±0.01)	0.62	(±0.04)	0.94	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Negative Exponential	0.2	0.86	(±0.01)	0.48	(±0.01)	0.92	(±0.01)
		0.6	0.91	(±0.01)	0.66	(±0.02)	0.95	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
EYT	3-categories QTL	0.2	0.77	(±0.04)	0.53	(±0.07)	0.84	(±0.03)
		0.6	0.75	(±0.03)	0.50	(±0.06)	0.83	(±0.02)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Uniform	0.2	0.75	(±0.03)	0.50	(±0.06)	0.83	(±0.02)
		0.6	0.80	(±0.03)	0.60	(±0.06)	0.87	(±0.02)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Negative Exponential	0.2	0.73	(±0.04)	0.47	(±0.07)	0.82	(±0.03)
		0.6	0.85	(±0.05)	0.70	(±0.09)	0.90	(±0.04)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)

Simulation experiment 4: Selections of crosses based on performance of lines in a previous cycle of variety development + MAS for SCN and PRR in early generations + phenotypic selection for yield in all field trial stages.

As with prior experiments, the analysis of variance of the simulated genotypic values for yield (Appendix A33) revealed that yield was affected by interactions involving combinations of genetic architecture and heritability with the founding sets of lines as well with stages of development. As with experiments 1, 2 and 3, the detection of significant interactions among the potential sources of variability for the simulated genotypic values was based on selection at each stage conducted on the phenotypic values of yield. As a consequence, simulated phenotypic yield on an entry mean basis of lines selected for evaluation in stages AYT-1, AYT-2 and EYT also were dependent upon the initial sets of crosses as well as lines selected from the PYT testing stages (Appendices A34-A36).

In experiment 4, analysis of variance revealed the simulated genotypic values for yield of EYT lines (Table 21) was affected by interactions with heritability, genetic architecture and founder sets of lines, as well with stages of development (Appendix A33).

Table 21: Analysis of variance for simulated genotypic values for yield due to founder sets, genetic architectures and heritability of EYT lines for a simulated experiment in which crosses consists of relative predictions from lines selected from AYTs and EYT in experiment 3, plus MAS for SCN and PRR traits plus phenotypic selection for yield.

Source	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	208919.8	8035.3	18552.6	<.0001*
Error	7004.7	0.4		
C. Total	215924.6			
Source		**S.Squares**	**F Ratio**	**Prob > F**
FounderSet		825.5	952.9	<.0001*
Genetic_architecture		74929.9	86501.6	<.0001*
FounderSet*Genetic_architecture		2460.0	1420.0	<.0001*
Heritability		113453.5	130974.7	<.0001*
FounderSet*Heritability		4256.2	2456.7	<.0001*
Genetic_architecture*Heritability		6455.5	3726.2	<.0001*
FounderSet*Genetic_architecture*Heritability		6538.9	1887.1	<.0001*

Although the specific lines were not consistent across all subsequent stages of evaluation, the distributions of the selected lines phenotypes were similar to those seen from actual field trials (Figure 10).

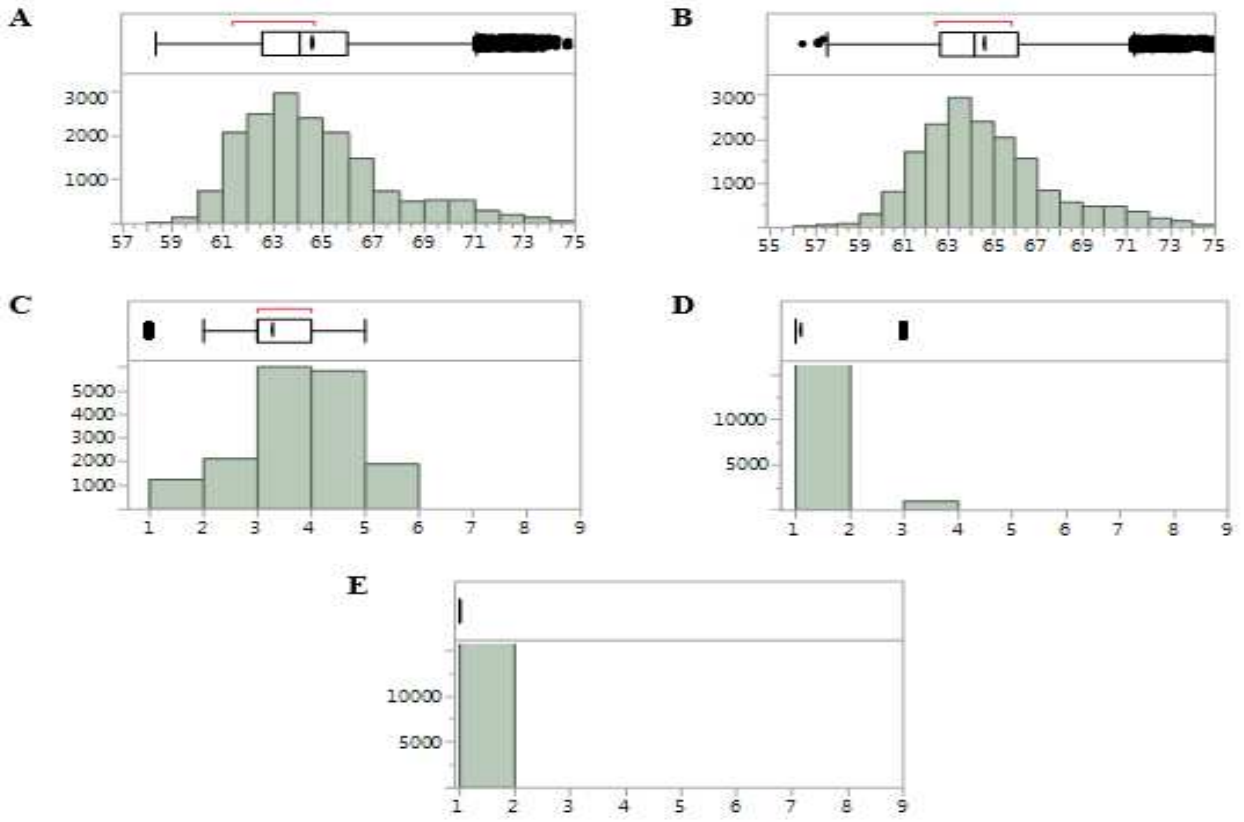

Figure 10: Distributions of simulated A) genotypic values of yield, B) phenotypic values of yield, C) lodging, D) PRR, and E) SCN of lines evaluated from experiment 4. Distributions are composed of data aggregated from lines selected for all combinations of simulated genetic architectures and heritabilities.

$F_{4:5}$ lines evaluated in AYT-1 had average simulated genotypic values for yield that ranged from 61.9-72.6 bu/ac, averaged phenotypic yields on an entry mean basis ranged from 62.0-73.1 bu/ac, lodging 2.1-3.9, SCN 1.0, and PRR 1.0-1.4. Subsequently, for each genetic architecture 40 AYT-2 $F_{4:7}$ lines (10 per RM group) exhibited simulated genotypic values for yield with values that ranged from 62.5-75.4, phenotypic yield ranged from 63.3-75.4 bu/ac, lodging 2.0-4.3, SCN 1.0, and PRR 1.0-1.6. Last, ten lines for evaluation as EYT $F_{4:7}$ lines had

the attributes to be released as new variety. Simulated genotypic values for yield ranged from 62.5-77.3 bu/ac, phenotypic yield 63.2-77.3 bu/ac, lodging 1.5-5.0, SCN 1.0, and PRR 1.0-2.0.

There were significant differences observed for simulated genotypic values for yield and averaged phenotypic yield among genetic architectures, stages of development and RM groups (Appendices A37-A39). Overall, genotypic and phenotypic yield values in AYT-2 and EYT were consistently higher than yield in AYT-1 (Table 22). There were no significant differences observed for lodging, SCN, and PRR among stages of development and RM groups.

Table 22: Genotypic values for yield and least square means for simulated yield, lodging, SCN and PRR in experiment 4. Values are listed by RM group and varietal development stages in a simulated experiment in which crosses consists of relative predictions from lines selected from AYTs and EYT in experiment 3, plus MAS for SCN and PRR traits plus phenotypic selection for yield.

Stage	RM	Genotypic yield mean	Phenotypic yield mean	Lodging mean	SCN mean	PRR mean
AYT-1	1.8-2.4	66.0 ^{A}	66.7 ^{A}	2.9 ^{A}	1.0 ^{A}	1.0 ^{A}
AYT-2	1.8-2.4	67.3 ^{B}	67.8 ^{B}	2.9 ^{A}	1.0 ^{A}	1.0 ^{A}
EYT	1.8-2.4	68.3 ^{B}	68.8 ^{B}	2.9 ^{A}	1.0 ^{A}	1.0 ^{A}
AYT-1	2.4-3.0	66.1 ^{A}	66.7 ^{A}	3.0 ^{A}	1.0 ^{A}	1.0 ^{A}
AYT-2	2.4-3.0	67.5 ^{B}	67.9 ^{B}	3.0 ^{A}	1.0 ^{A}	1.0 ^{A}
EYT	2.4-3.0	68.3 ^{B}	68.8 ^{B}	3.1 ^{A}	1.0 ^{A}	1.0 ^{A}
AYT-1	3.0-3.6	66.3 ^{A}	67.0 ^{A}	3.0 ^{A}	1.0 ^{A}	1.0 ^{A}
AYT-2	3.0-3.6	67.7 ^{B}	68.0 ^{B}	3.1 ^{A}	1.0 ^{A}	1.1 ^{A}
EYT	3.0-3.6	68.3 ^{B}	68.7 ^{B}	3.0 ^{A}	1.0 ^{A}	1.1 ^{A}
AYT-1	3.6-4.2	65.9 ^{A}	66.6 ^{A}	3.2 ^{A}	1.0 ^{A}	1.1 ^{A}
AYT-2	3.6-4.2	67.3 ^{B}	67.7 ^{B}	3.2 ^{A}	1.0 ^{A}	1.0 ^{A}
EYT	3.6-4.2	67.7 ^{B}	68.1 ^{B}	3.1 ^{A}	1.0 ^{A}	1.0 ^{A}

Levels not connected by the same letter are significantly different (Tukey α=0.01)

Differences among decision accuracy, sensitivity and specificity of selected lines for yield are reported for all three sets of founder lines in which genetic architectures explain 20% and 60% of the phenotypic variance on an entry mean basis (Appendices A40-A42). Results are similar to previous experiments, revealing high values for decision accuracy, specificity, and sensitivity showing the dependency of heritability and genetic architectures combinations (Table 23).

Table 23: Average decision accuracy, sensitivity, and specificity metrics from three founding sets of crosses used to initiate variety development for each heritability and genetic architecture combination and stage of development for experiment 4. The three sets of crosses were selected based on predicted progeny values, where the training set consisted of related lines selected from AYTs and EYT in experiment 3, plus MAS for SCN and PRR traits plus phenotypic selection for yield.

Stage	Yield genetic architecture	H^2	Accuracy	(SE)	Sensitivity	(SE)	Specificity	(SE)
PYT	3-categories QTL	0.2	0.87	(±0.01)	0.35	(±0.01)	0.93	(±0.01)
		0.6	0.93	(±0.01)	0.65	(±0.01)	0.96	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Uniform	0.2	0.88	(±0.01)	0.40	(±0.02)	0.93	(±0.01)
		0.6	0.94	(±0.01)	0.68	(±0.03)	0.96	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Negative Exponential	0.2	0.87	(±0.01)	0.36	(±0.02)	0.93	(±0.01)
		0.6	0.93	(±0.01)	0.65	(±0.03)	0.96	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
AYT-1	3-categories QTL	0.2	0.80	(±0.01)	0.48	(±0.01)	0.87	(±0.01)
		0.6	0.86	(±0.01)	0.63	(±0.03)	0.91	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Uniform	0.2	0.83	(±0.01)	0.57	(±0.01)	0.90	(±0.01)
		0.6	0.88	(±0.01)	0.70	(±0.01)	0.93	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Negative Exponential	0.2	0.81	(±0.01)	0.52	(±0.02)	0.88	(±0.01)
		0.6	0.88	(±0.02)	0.69	(±0.03)	0.92	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
AYT-2	3-categories QTL	0.2	0.86	(±0.01)	0.48	(±0.02)	0.92	(±0.01)
		0.6	0.90	(±0.01)	0.63	(±0.01)	0.94	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Uniform	0.2	0.86	(±0.01)	0.46	(±0.04)	0.92	(±0.01)
		0.6	0.89	(±0.01)	0.60	(±0.03)	0.94	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Negative Exponential	0.2	0.86	(±0.01)	0.49	(±0.04)	0.92	(±0.01)
		0.6	0.92	(±0.01)	0.69	(±0.03)	0.95	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
EYT	3-categories QTL	0.2	0.72	(±0.02)	0.43	(±0.04)	0.81	(±0.02)
		0.6	0.83	(±0.05)	0.67	(±0.09)	0.89	(±0.03)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Uniform	0.2	0.83	(±0.02)	0.67	(±0.04)	0.89	(±0.02)
		0.6	0.83	(±0.04)	0.67	(±0.07)	0.89	(±0.03)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Negative Exponential	0.2	0.70	(±0.06)	0.40	(±0.10)	0.80	(±0.04)
		0.6	0.83	(±0.04)	0.67	(±0.07)	0.89	(±0.03)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)

Simulation experiment 5: Selection of crosses based on performance of relatives in the previous cycle of variety development + MAS for SCN and PRR in early generation + genomic selection in the PYT from the same training set used for selection of crosses

Three sets of crosses used to initiate variety development in experiment 5 were the same as the sets used in experiment 4. Accuracy (Pearson's correlation between simulated genotypic values and simulated phenotypic values for lines in the PYT) was calculated from all nine combinations of heritabilities nested within yield genetic architectures for all three founder sets. Estimates of accuracy for the first set of founder lines ranged between 0.48-0.52 (± 0.03), 0.62-0.67 (± 0.03), and 0.59-0.78 (± 0.02) for genetic architectures that explained 20%, 60%, and 100% of the phenotypic variance on an entry mean basis, respectively (Figure 11). For the second set of founder lines, estimates of accuracy ranged between 0.48-0.49 (± 0.04), 0.57-0.62 (± 0.02), and 0.59-0.84 (± 0.02). For the third set of founder lines, estimates of accuracy ranged from 0.44-0.53 (± 0.03), 0.57-0.65 (± 0.03), and 0.62-0.80 (± 0.02).

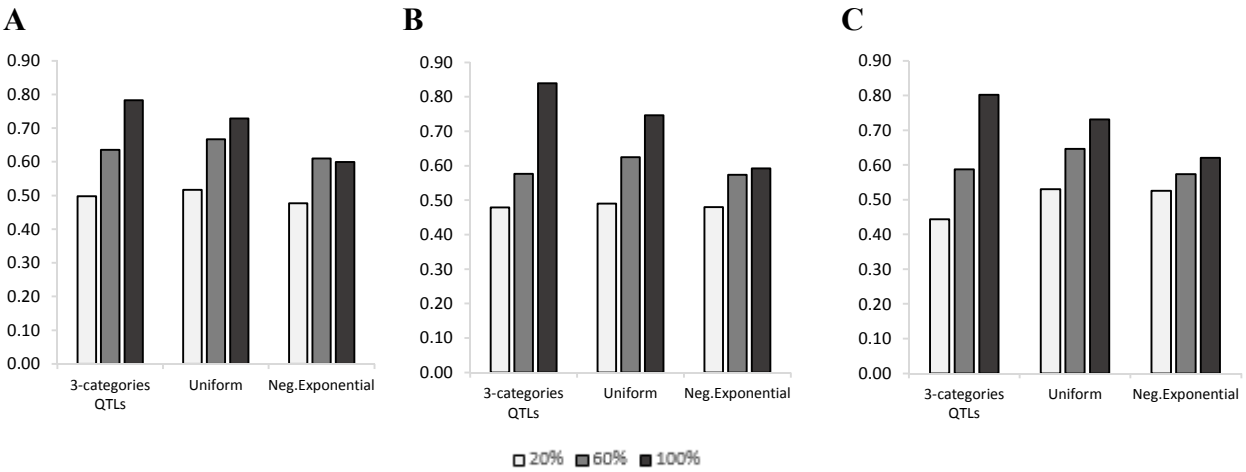

Figure 11: Accuracy (left axis) measured as Pearson`s correlation. The genetic architecture values correspond to 20%, 60%, and 100% of the phenotypic variance on an entry mean basis. Training sets consisted of simulated genotypic and phenotypic data from 1,540 lines evaluated in AYT-1 and 300 lines evaluated in AYT-2 in combination with genotypic assays at 1,000 loci. A) First set of founder lines, B) Second set of founder lines, C) Third set of founder lines.

Analysis of variance of the simulated genotypic values for yield (Appendix 43) in experiment 5 revealed large interactions involving combinations of genetic architecture and

heritability with the founding sets of lines as well with stages of development. These results are consistent with other variety development experiments. Detection of significant interactions among the potential sources of variability for the simulated genotypic values was also reflected for phenotypic yields, on an entry mean basis. As a consequence, phenotypic yield data for lines selected for evaluation stages (AYT-1, AYT-2 and EYT) were dependent upon the initial sets of lines selected from the PYT testing stages (Appendices A44-A46). Although the specific lines were not consistent across all subsequent stages of evaluation, the distribution of the phenotypes of selected lines were similar to those seen from actual field trials (data not shown).

In experiment 5, analysis of variance revealed the simulated genotypic values for yield of EYT lines (Table 24) were affected by interactions with heritability, genetic architecture and founder sets of lines as well with stages of development (Appendix A43).

Table 24: Analysis of variance of average genotypic values for yield for EYT lines showing significant interaction between founder set, genetic architecture and heritability. Experiment 5 consists of relative predictions from lines selected from AYTs and EYT in experiment 3, plus MAS for SCN and PRR traits plus genotypic selection for yield using GEBVs for PYT lines.

Source	**Sum of Squares**	**Mean Square**	**F Ratio**	**Prob > F**
Model	206468.2	7941.0	19493.1	<.0001*
Error	6588.5	0.4		
C. Total	213056.8			
Source		**S.Squares**	**F Ratio**	**Prob > F**
FounderSet		721.3	885.3	<.0001*
GeneticArchitecture		73759.5	90529.7	<.0001*
FounderSet*GeneticArchitecture		2619.3	1607.4	<.0001*
Heritability		112008.3	137474.9	<.0001*
FounderSet*Heritability		4354.1	2672.0	<.0001*
GeneticArchitecture*Heritability		6931.9	4253.9	<.0001*
FounderSet*GeneticArchitecture*Heritability		6073.7	1863.6	<.0001*

$F_{4:5}$ lines evaluated in AYT-1 had average simulated genotypic values for yield that ranged from 62.0-72.6 bu/ac, averaged phenotypic yield on an entry mean basis ranged from 63.0-72.6 bu/ac, lodging 2.1-4.0, SCN 1.0, and PRR 1.0-1.3. Then, for each genetic architecture 40 AYT-2 $F_{4:7}$ lines (10 by RM group) were selected for EYT with simulated genotypic yield

range from 62.4-75.4 bu/ac, phenotypic yield range from 63.2-75.4 bu/ac, lodging 2.0-4.3, SCN 1.0, and PRR 1.0-1.6. Last, ten high yield EYT $F_{4:7}$ lines had the attributes to be released as new variety. Simulated genotypic values for yield ranged from 62.6-77.3 bu/ac, averaged phenotypic yield ranged from 63.4-77.3 bu/ac, lodging 1.0-5.0, SCN 1.0, and PRR 1.0-2.0.

There were significant differences observed for simulated genotypic values for yield and averaged simulated phenotypic values for yield (Appendices A47-A49). There were no significant differences for simulated lodging, SCN and PRR values (Table 25).

Table 25: Genotypic values and least square means for simulated yield, lodging, SCN and PRR in experiment 5. Values are listed by RM group and varietal development stages in a simulated experiment that consists of relative predictions from lines selected from AYTs and EYT in experiment 3, plus MAS for SCN and PRR traits plus genotypic selection for yield using GEBVs for PYT lines.

Stage	RM	Genotypic yield mean	Phenotypic yield mean	Lodging mean	SCN mean	PRR mean
AYT-1	1.8-2.4	66.0 ^{A}	66.7 ^{A}	2.9 ^{A}	1.0 ^{A}	1.0 ^{A}
AYT-2	1.8-2.4	67.3 ^{B}	67.8 ^{B}	2.9 ^{A}	1.0 ^{A}	1.0 ^{A}
EYT	1.8-2.4	68.4 ^{B}	68.9 ^{C}	2.9 ^{A}	1.0 ^{A}	1.0 ^{A}
AYT-1	2.4-3.0	66.1 ^{A}	66.8 ^{A}	3.0 ^{A}	1.0 ^{A}	1.0 ^{A}
AYT-2	2.4-3.0	67.5 ^{B}	68.0 ^{B}	3.0 ^{A}	1.0 ^{A}	1.0 ^{A}
EYT	2.4-3.0	68.4 ^{B}	68.8 ^{B}	3.0 ^{A}	1.0 ^{A}	1.0 ^{A}
AYT-1	3.0-3.6	66.4 ^{A}	67.0 ^{A}	3.0 ^{A}	1.0 ^{A}	1.0 ^{A}
AYT-2	3.0-3.6	67.7 ^{B}	68.1 ^{B}	3.1 ^{A}	1.0 ^{A}	1.1 ^{A}
EYT	3.0-3.6	68.3 ^{B}	68.7 ^{B}	3.1 ^{A}	1.0 ^{A}	1.0 ^{A}
AYT-1	3.6-4.2	66.0 ^{A}	66.7 ^{A}	3.2 ^{A}	1.0 ^{A}	1.1 ^{A}
AYT-2	3.6-4.2	67.3 ^{B}	67.7 ^{B}	3.1 ^{A}	1.0 ^{A}	1.1 ^{A}
EYT	3.6-4.2	67.6 ^{B}	68.0 ^{B}	3.1 ^{A}	1.0 ^{A}	1.1 ^{A}

Levels not connected by the same letter are significantly different (Tukey α=0.01)

Decision accuracy, sensitivity, and specificity metrics were calculated for the selection sets (Appendices A50-A52). Different from experiments 1 to 4, decision accuracy, sensitivity and specificity for PYTs due to selections based on GEBVs were not equal to 1.00 when all of the simulated phenotypic variability is explained by simulated genotypic variability (Table 26). On the other hand, decision accuracy, specificity, and sensitivity from AYT-1 to EYT are comparable to results described in experiments 1, 2, 3, and 4.

Table 26: Average decision accuracy, sensitivity, and specificity from three founding sets of crosses used to initiate variety development for nine combinations of genetic architecture and heritability and stage of development in soybean variety development for experiment 5 in which crosses consists of relative predictions from lines selected from AYTs and EYT in experiment 3, plus MAS for SCN and PRR traits plus phenotypic selection for yield.

Stage	Yield genetic architecture	H^2	Accuracy	(SE)	Sensitivity	(SE)	Specificity	(SE)
PYT	3-categories QTL	0.2	0.89	(±0.02)	0.44	(±0.09)	0.94	(±0.01)
		0.6	0.91	(±0.01)	0.55	(±0.05)	0.95	(±0.01)
		1.0	0.94	(±0.01)	0.70	(±0.02)	0.97	(±0.01)
	Uniform	0.2	0.89	(±0.02)	0.47	(±0.10)	0.94	(±0.02)
		0.6	0.92	(±0.02)	0.62	(±0.09)	0.96	(±0.01)
		1.0	0.94	(±0.01)	0.69	(±0.01)	0.96	(±0.01)
	Negative Exponential	0.2	0.88	(±0.02)	0.38	(±0.06)	0.93	(±0.01)
		0.6	0.91	(±0.02)	0.55	(±0.06)	0.95	(±0.01)
		1.0	0.92	(±0.01)	0.59	(±0.03)	0.95	(±0.01)
AYT-1	3-categories QTL	0.2	0.79	(±0.01)	0.47	(±0.02)	0.87	(±0.01)
		0.6	0.87	(±0.01)	0.67	(±0.02)	0.92	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Uniform	0.2	0.81	(±0.01)	0.50	(±0.01)	0.88	(±0.01)
		0.6	0.88	(±0.01)	0.70	(±0.01)	0.93	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Negative Exponential	0.2	0.80	(±0.01)	0.47	(±0.02)	0.87	(±0.01)
		0.6	0.89	(±0.01)	0.71	(±0.02)	0.93	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
AYT-2	3-categories QTL	0.2	0.86	(±0.01)	0.46	(±0.03)	0.92	(±0.01)
		0.6	0.90	(±0.01)	0.64	(±0.01)	0.94	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Uniform	0.2	0.88	(±0.01)	0.53	(±0.03)	0.93	(±0.01)
		0.6	0.90	(±0.01)	0.61	(±0.04)	0.94	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Negative Exponential	0.2	0.87	(±0.01)	0.50	(±0.03)	0.92	(±0.01)
		0.6	0.92	(±0.01)	0.69	(±0.03)	0.95	(±0.01)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
EYT	3-categories QTL	0.2	0.72	(±0.02)	0.43	(±0.04)	0.81	(±0.02)
		0.6	0.85	(±0.05)	0.70	(±0.10)	0.90	(±0.04)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Uniform	0.2	0.83	(±0.02)	0.67	(±0.04)	0.89	(±0.02)
		0.6	0.82	(±0.02)	0.63	(±0.04)	0.88	(±0.02)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)
	Negative Exponential	0.2	0.73	(±0.05)	0.47	(±0.09)	0.82	(±0.03)
		0.6	0.83	(±0.04)	0.67	(±0.07)	0.89	(±0.03)
		1.0	1.00	(±0.00)	1.00	(±0.00)	1.00	(±0.00)

Comparison of experiments 1 and 2: Impact of MAS in early generations for non-yield traits

Combined analysis of variance of the EYT simulated genotypic values for yield for experiments 1 and 2 (Table 27) revealed that both experiments are influenced by the interactions between heritability and genetic architecture with experiment.

Table 27: Combined analysis of variance for EYT simulated genotypic values for yield due to founder sets, genetic architectures and heritability of lines for experiment 1 with only phenotypic selection for yield, and experiment 2 simulated with MAS for SCN and PRR traits + phenotypic selection for yield.

Source	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	2505.8	47.2	91.3	<.0001*
Error	251.4	0.5		
C. Total	2757.3			

Source	S. Squares	F Ratio	Prob > F
Experiment	79.9	154.5	<.0001*
Founder_Set	5.0	4.8	0.0083*
Experiment*Founder_Set	0.6	0.6	0.5248
Genetic_Architecture	958.3	926.0	<.0001*
Experiment*Genetic_Architecture	25.9	25.0	<.0001*
Founder_Set*Genetic_Architecture	3.1	1.5	0.1909
Experiment*Founder_Set*Genetic_Architecture	2.6	1.2	0.2750
Heritability	1333.3	1288.3	<.0001*
Experiment*Heritability	7.0	6.8	0.0012*
Founder_Set*Heritability	8.9	4.3	0.0020*
Experiment*Founder_Set*Heritability	0.8	0.4	0.7938
Genetic_Architecture*Heritability	11.6	5.6	0.0002*
Experiment*Genetic_Architecture*Heritability	61.7	29.8	<.0001*
Founder_Set*Genetic_Architecture*Heritability	4.7	1.1	0.3366
Experiment*Founder_Set*Genetic_Arch*Heritability	1.6	0.4	0.9183

Trait means comparison between experiments 1 and 2, revealed significant differences for simulated genotypic values for yield and simulated phenotypic values for yield (Table 28). However, there were no significant differences between experiment 1 and 2 for yield from the EYT stage, nor were there significant differences for simulated lodging (Tukey α=0.01). Significant differences were not observed for SCN, except for AYT-1 in experiment 1 in all four

RM groups. Significant differences between experiment 1 and 2were observed for PRR from the EYT stage.

Table 28: Least Square Means for simulated genotypic values of yield, phenotypic yield, lodging, SCN and PRR. Comparisons are conducted by RM group and varietal development stages for experiments 1 and 2.

Stage_Experiment	RM	Genotypic yield mean	Phenotypic yield mean	Lodging mean	SCN mean	PRR mean
AYT-1_Exp1	1.8-2.4	63.6 ^{A}	64.3 ^{A}	3.0 ^{A}	5.0 ^{A}	3.7 ^{A}
AYT-1_Exp2		63.6 ^{A}	64.3 ^{A}	3.1 ^{A}	1.0 ^{B}	1.1 ^{C}
AYT-2_Exp1		64.0 ^{B}	64.1 ^{A}	3.1 ^{A}	1.0 ^{B}	1.7 ^{B}
AYT-2_Exp2		64.8 ^{C}	65.3 ^{B}	3.1 ^{A}	1.0 ^{B}	1.1 ^{C}
EYT_Exp1		64.7 ^{C}	65.2 ^{B}	3.3 ^{A}	1.0 ^{B}	1.6 ^{B}
EYT_Exp2		65.4 ^{C}	66.0 ^{B}	3.1 ^{A}	1.0 ^{B}	1.1 ^{C}
AYT-1_Exp1	2.4-3.0	64.0 ^{A}	64.7 ^{B}	3.1 ^{A}	4.8 ^{A}	3.9 ^{A}
AYT-1_Exp2		63.8 ^{A}	64.5 ^{A}	3.1 ^{A}	1.0 ^{B}	1.1 ^{C}
AYT-2_Exp1		64.3 ^{B}	64.6 ^{A}	3.2 ^{A}	1.0 ^{B}	1.7 ^{B}
AYT-2_Exp2		65.1 ^{C}	65.4 ^{C}	3.2 ^{A}	1.0 ^{B}	1.1 ^{C}
EYT_Exp1		65.0 ^{C}	65.4 ^{C}	3.3 ^{A}	1.0 ^{B}	1.6 ^{B}
EYT_Exp2		65.6 ^{C}	66.0 ^{C}	3.2 ^{A}	1.0 ^{B}	1.1 ^{C}
AYT-1_Exp1	3.0-3.6	64.1 ^{A}	64.9 ^{A}	3.1 ^{A}	4.9 ^{A}	3.8 ^{A}
AYT-1_Exp2		64.0 ^{A}	64.7 ^{A}	3.1 ^{A}	1.0 ^{B}	1.1 ^{C}
AYT-2_Exp1		64.5 ^{B}	64.7 ^{A}	3.3 ^{A}	1.0 ^{B}	1.7 ^{B}
AYT-2_Exp2		65.3 ^{C}	65.7 ^{B}	3.2 ^{A}	1.0 ^{B}	1.1 ^{C}
EYT_Exp1		65.2 ^{C}	65.6 ^{B}	3.5 ^{A}	1.0 ^{B}	1.5 ^{B}
EYT_Exp2		66.0 ^{C}	66.6 ^{C}	3.4 ^{A}	1.0 ^{B}	1.1 ^{C}
AYT-1_Exp1	3.6-4.2	64.1 ^{A}	64.8 ^{A}	3.2 ^{A}	4.6 ^{A}	3.7 ^{A}
AYT-1_Exp2		64.0 ^{A}	64.7 ^{A}	3.2 ^{A}	1.0 ^{B}	1.1 ^{C}
AYT-2_Exp1		64.6 ^{B}	64.8 ^{A}	3.2 ^{A}	1.0 ^{B}	1.6 ^{B}
AYT-2_Exp2		65.4 ^{C}	65.7 ^{B}	3.2 ^{A}	1.0 ^{B}	1.1 ^{C}
EYT_Exp1		65.3 ^{C}	65.8 ^{C}	3.2 ^{A}	1.0 ^{B}	1.6 ^{B}
EYT_Exp2		66.3 ^{C}	66.6 ^{C}	3.1 ^{A}	1.0 ^{B}	1.1 ^{C}

Levels not connected by the same letter are significantly different (Tukey α=0.01)

Comparison of experiments 2 and 3: Impact of using GEBV's to select founding crosses based on performance of founders as the training set

Combined analysis of variance of the EYT simulated genotypic values for yield for experiments 2 and 3 (Table 29) revealed that both experiments are influenced by interactions between heritability and genetic architecture with experiment.

Table 29: Combined analysis of variance for EYT simulated genotypic values for yield due to founder sets, genetic architectures and heritability of lines for experiment 2 and 3 simulated with MAS for SCN and PRR traits + phenotypic selection for yield.

Source	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	2505.8	47.2	91.3	<.0001*
Error	251.4	0.5		
C. Total	2757.3			

Source	S. Squares	F Ratio	Prob > F
Experiment	0.7	1.4	0.2336
Founder_Set	1.6	1.5	0.2184
Experiment*Founder_Set	3.7	3.4	0.0312*
Genetic_Architecture	1217.6	1140.1	<.0001*
Experiment*Genetic_Architecture	1.9	1.7	0.1694
Founder_Set*Genetic_Architecture	5.1	2.4	0.0470*
Experiment*Founder_Set*Genetic_Architecture	1.7	0.8	0.5129
Heritability	1609.2	1506.8	<.0001*
Experiment*Heritability	0.9	0.8	0.4180
Founder_Set*Heritability	2.7	1.2	0.2824
Experiment*Founder_Set*Heritability	3.9	1.8	0.1148
Genetic_Architecture*Heritability	2.0	0.9	0.4340
Experiment*Genetic_Architecture*Heritability	46.1	21.6	<.0001*
Founder_Set*Genetic_Architecture*Heritability	1.2	0.2	0.9712
Experiment*Founder_Set*Genetic_Arch*Heritability	5.5	1.3	0.2398

Trait means comparison between experiments 2 and 3, revealed significant differences for simulated genotypic values for yield and simulated phenotypic values for yield between AYT-1 with AYT-2 and EYT. However, significant differences were not observed between AYT-2 with EYT (Table 30). There were no significant differences for simulated lodging, SCN and PRR (Tukey α=0.01).

Table 30: Least Square Means for simulated genotypic values of yield, phenotypic yield, lodging, SCN and PRR. Comparisons are conducted by RM group and varietal development stages for experiments 2 and 3.

Stage_Experiment	RM	Genotypic yield mean	Phenotypic yield mean	Lodging mean	SCN mean	PRR mean
AYT-1_Exp2	1.8-2.4	63.6 A	64.3 A	3.1 A	1.0 A	1.1 A
AYT-1_Exp3		63.7 A	64.4 A	3.1 A	1.0 A	1.2 A
AYT-2_Exp2		64.8 B	65.3 B	3.1 A	1.0 A	1.1 A
AYT-2_Exp3		64.9 B	65.4 B	3.1 A	1.0 A	1.1 A
EYT_Exp2		65.4 B	66.0 B	3.1 A	1.0 A	1.1 A
EYT_Exp3		65.4 B	66.0 B	2.9 A	1.0 A	1.1 A
AYT-1_Exp2	2.4-3.0	63.8 A	64.5 A	3.1 A	1.0 A	1.2 A
AYT-1_Exp3		63.9 A	64.7 B	3.1 A	1.0 A	1.1 A
AYT-2_Exp2		65.1 B	65.4 C	3.1 A	1.0 A	1.1 A
AYT-2_Exp3		65.2 B	65.7 C	3.1 A	1.0 A	1.2 A
EYT_Exp2		65.6 B	66.0 D	3.2 A	1.0 A	1.1 A
EYT_Exp3		65.7 B	66.2 D	3.2 A	1.0 A	1.1 A
AYT-1_Exp2	3.0-3.6	64.0 A	64.7 A	3.1 A	1.0 A	1.1 A
AYT-1_Exp3		64.3 B	65.0 B	3.1 A	1.0 A	1.2 A
AYT-2_Exp2		65.3 C	65.7 C	3.2 A	1.0 A	1.1 A
AYT-2_Exp3		65.5 C	65.9 C	3.0 A	1.0 A	1.2 A
EYT_Exp2		66.0 D	66.6 D	3.4 A	1.0 A	1.1 A
EYT_Exp3		66.2 D	66.5 D	3.0 A	1.0 A	1.2 A
AYT-1_Exp2	3.6-4.2	64.0 A	64.7 A	3.2 A	1.0 A	1.1 A
AYT-1_Exp3		64.2 A	64.9 B	3.2 A	1.0 A	1.2 A
AYT-2_Exp2		65.4 B	65.7 C	3.2 A	1.0 A	1.1 A
AYT-2_Exp3		65.5 B	65.9 C	3.2 A	1.0 A	1.2 A
EYT_Exp2		66.3 B	66.6 D	3.1 A	1.0 A	1.1 A
EYT_Exp3		66.2 B	66.7 D	3.1 A	1.0 A	1.2 A

Levels not connected by the same letter are significantly different (Tukey α=0.01)

Comparison of experiments 3 and 4: Impact of using founding crosses based on performance of relatives from a previous variety development project

Combined analysis of variance of the EYT simulated genotypic values for yield for experiments 3 and 4 (Table 31) revealed that both experiments are influenced by interactions existing between heritability and genetic architecture with either experiment and founder set.

Table 31: Combined analysis of variance for EYT simulated genotypic values for yield due to founder sets, genetic architectures and heritability of lines for experiment 3 and 4 simulated with founder and relative prediction plus MAS for SCN and PRR traits plus phenotypic selection for yield.

Source	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	5621.8	106.0	91.3	<.0001*
Error	246.8	0.5		
C. Total	5868.7			
Source		**S. Squares**	**F Ratio**	**Prob > F**
Experiment		692.6	1363.4	<.0001*
Founder_Set		12.6	12.4	<.0001*
Experiment*Founder_Set		2.9	2.9	0.0552
Genetic_Architecture		1754.9	1727.2	<.0001*
Experiment*Genetic_Architecture		66.9	65.9	<.0001*
Founder_Set*Genetic_Architecture		24.7	12.1	<.0001*
Experiment*Founder_Set*Genetic_Architecture		18.6	9.1	<.0001*
Heritability		2606.1	2565.0	<.0001*
Experiment*Heritability		127.2	125.2	<.0001*
Founder_Set*Heritability		38.0	18.7	<.0001*
Experiment*Founder_Set*Heritability		33.5	16.5	<.0001*
Genetic_Architecture*Heritability		100.5	49.4	<.0001*
Experiment*Genetic_Architecture*Heritability		30.8	15.1	<.0001*
Founder_Set*Genetic_Architecture*Heritability		61.0	15.0	<.0001*
Experiment*Founder_Set*Genetic_Arch*Heritability		50.9	12.5	<.0001*

Significant differences were observed for simulated genotypic values for yield and simulated phenotypic values for yield between AYT-1 with AYT-2 and EYT values (Table 32). Overall, significant differences were observed for all four RM groups of simulated EYT genotypic and phenotypic values for yield. Notably the final outcomes for genotypic values for yield from the EYT stage were about 5% greater than they were from experiment 3, where the training sets were not based on performance of relatives. There were no significant differences for simulated lodging, SCN and PRR (Tukey α=0.01).

Table 32: Least Square Means for simulated genotypic values of yield, phenotypic yield, lodging, SCN and PRR. Comparisons are conducted by RM group and varietal development stages for experiments 3 and 4.

Stage_Experiment	RM	Genotypic yield mean	Phenotypic yield mean	Lodging mean	SCN mean	PRR mean
AYT-1_Exp3	1.8-2.4	63.7 ^{A}	64.4 ^{A}	3.1 ^{A}	1.0 ^{A}	1.2 ^{A}
AYT-1_Exp4		66.0 ^{C}	66.7 ^{C}	2.9 ^{A}	1.0 ^{A}	1.0 ^{A}
AYT-2_Exp3		64.9 ^{B}	65.4 ^{B}	3.1 ^{A}	1.0 ^{A}	1.1 ^{A}
AYT-2_Exp4		67.3 ^{D}	67.8 ^{D}	2.9 ^{A}	1.0 ^{A}	1.0 ^{A}
EYT_Exp3		65.4 ^{C}	66.0 ^{B}	2.9 ^{A}	1.0 ^{A}	1.1 ^{A}
EYT_Exp4		68.2 ^{D}	68.8 ^{E}	2.9 ^{A}	1.0 ^{A}	1.0 ^{A}
AYT-1_Exp3	2.4-3.0	63.9 ^{A}	64.7 ^{A}	3.1 ^{A}	1.0 ^{A}	1.1 ^{A}
AYT-1_Exp4		66.0 ^{C}	66.7 ^{C}	3.0 ^{A}	1.0 ^{A}	1.0 ^{A}
AYT-2_Exp3		65.2 ^{B}	65.7 ^{B}	3.1 ^{A}	1.0 ^{A}	1.2 ^{A}
AYT-2_Exp4		67.5 ^{D}	67.9 ^{D}	3.0 ^{A}	1.0 ^{A}	1.0 ^{A}
EYT_Exp3		65.7 ^{C}	66.2 ^{C}	3.2 ^{A}	1.0 ^{A}	1.1 ^{A}
EYT_Exp4		68.3 ^{D}	68.8 ^{D}	3.1 ^{A}	1.0 ^{A}	1.0 ^{A}
AYT-1_Exp3	3.0-3.6	64.3 ^{A}	65.0 ^{A}	3.1 ^{A}	1.0 ^{A}	1.2 ^{A}
AYT-1_Exp4		66.3 ^{C}	67.0 ^{C}	3.1 ^{A}	1.0 ^{A}	1.0 ^{A}
AYT-2_Exp3		65.5 ^{B}	65.9 ^{B}	3.0 ^{A}	1.0 ^{A}	1.2 ^{A}
AYT-2_Exp4		67.6 ^{D}	68.1 ^{D}	3.1 ^{A}	1.0 ^{A}	1.1 ^{A}
EYT_Exp3		66.2 ^{C}	66.5 ^{C}	3.0 ^{A}	1.0 ^{A}	1.2 ^{A}
EYT_Exp4		68.3 ^{D}	68.7 ^{D}	3.0 ^{A}	1.0 ^{A}	1.1 ^{A}
AYT-1_Exp3	3.6-4.2	64.2 ^{A}	64.9 ^{A}	3.2 ^{A}	1.0 ^{A}	1.2 ^{A}
AYT-1_Exp4		65.9 ^{B}	66.6 ^{C}	3.2 ^{A}	1.0 ^{A}	1.1 ^{A}
AYT-2_Exp3		65.5 ^{B}	65.9 ^{B}	3.2 ^{A}	1.0 ^{A}	1.2 ^{A}
AYT-2_Exp4		67.3 ^{C}	67.7 ^{D}	3.2 ^{A}	1.0 ^{A}	1.1 ^{A}
EYT_Exp3		66.2 ^{B}	66.7 ^{C}	3.1 ^{A}	1.0 ^{A}	1.3 ^{A}
EYT_Exp4		67.6 ^{C}	68.2 ^{D}	3.1 ^{A}	1.0 ^{A}	1.0 ^{A}

Levels not connected by the same letter are significantly different (Tukey α=0.01)

Comparison of experiments 4 and 5: Impact of using GEBV's to select lines in a PYT, when GEBVs are based on performance of relatives' in a previous variety development project

Because the sets of lines that entered into the PYT were the same for experiment 4 and 5, it was possible to evaluate whether phenotypic yield means from the phenotypic and GEBV-derived set of lines were the same (Table 33). Combined analysis of variance of the EYT simulated genotypic values for yield for experiments 4 and 5 shows that heritability and genetic architecture combinations interact with founder sets.

Table 33: Combined analysis of variance for EYT simulated genotypic values for yield due to founder sets, genetic architectures and heritability of lines for experiment 3 and 4 simulated with relative predictions plus MAS for SCN and PRR traits plus phenotypic and GEBV selection for yield.

Source	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	6922.2	130.6	280.3	<.0001*
Error	226.4	0.4		
C. Total	7148.7			
Source		**S. Squares**	**F Ratio**	**Prob > F**
Experiment		0.0	0.0	0.8878
Founder_Set		25.7	27.6	<.0001*
Experiment*Founder_Set		0.0	0.0	0.9344
Genetic_Architecture		2477.2	2658.5	<.0001*
Experiment*Genetic_Architecture		0.0	0.0	0.9558
Founder_Set*Genetic_Architecture		84.6	45.4	<.0001*
Experiment*Founder_Set*Genetic_Architecture		0.0	0.0	0.9992
Heritability		3422.3	3672.8	<.0001*
Experiment*Heritability		0.0	0.0	0.9145
Founder_Set*Heritability		478.0	256.4	<.0001*
Experiment*Founder_Set*Heritability		0.3	0.2	0.9340
Genetic_Architecture*Heritability		396.4	212.7	<.0001*
Experiment*Genetic_Architecture*Heritability		0.1	0.0	0.9878
Founder_Set*Genetic_Architecture*Heritability		36.4	9.7	<.0001*
Experiment*Founder_Set*Genetic_Arch*Heritability		0.5	0.1	0.9969

Significant differences were observed for simulated genotypic values for yield and simulated phenotypic values for yield between AYT-1 with AYT-2 and EYT. Significant differences were not observed between EYT in experiment 4 and 5 (Table 34). There were no significant differences for simulated lodging, SCN and PRR (Tukey α=0.01).

Table 34: Least Square Means for simulated genotypic values of yield, phenotypic yield, lodging, SCN and PRR. Comparisons are conducted by RM group and varietal development stages for experiments 4 and 5.

Stage_Experiment	RM	Genotypic yield mean	Phenotypic yield mean	Lodging mean	SCN mean	PRR mean
AYT-1_Exp4	1.8-2.4	66.0 A	66.7 A	2.9 A	1.0 A	1.0 A
AYT-1_Exp5		66.0 A	66.7 A	2.9 A	1.0 A	1.0 A
AYT-2_Exp4		67.3 B	67.8 B	2.9 A	1.0 A	1.0 A
AYT-2_Exp5		67.3 B	67.8 B	2.9 A	1.0 A	1.0 A
EYT_Exp4		68.2 B	68.8 B	2.9 A	1.0 A	1.0 A
EYT_Exp5		68.4 B	68.9 B	2.9 A	1.0 A	1.0 A
AYT-1_Exp4	2.4-3.0	66.0 A	66.7 A	3.0 A	1.0 A	1.0 A
AYT-1_Exp5		66.1 A	66.8 A	3.0 A	1.0 A	1.0 A
AYT-2_Exp4		67.5 B	67.9 B	3.0 A	1.0 A	1.0 A
AYT-2_Exp5		67.5 B	67.9 B	3.0 A	1.0 A	1.0 A
EYT_Exp4		68.3 B	68.8 B	3.1 A	1.0 A	1.0 A
EYT_Exp5		68.4 B	68.8 B	3.0 A	1.0 A	1.0 A
AYT-1_Exp4	3.0-3.6	66.3 A	67.0 A	3.0 A	1.0 A	1.0 A
AYT-1_Exp5		66.3 A	67.0 A	3.0 A	1.0 A	1.0 A
AYT-2_Exp4		67.6 B	68.1 B	3.1 A	1.0 A	1.1 A
AYT-2_Exp5		67.7 B	68.1 B	3.1 A	1.0 A	1.0 A
EYT_Exp4		68.3 B	68.7 B	3.0 A	1.0 A	1.1 A
EYT_Exp5		68.3 B	68.7 B	3.1 A	1.0 A	1.0 A
AYT-1_Exp4	3.6-4.2	65.9 A	66.6 A	3.2 A	1.0 A	1.1 A
AYT-1_Exp5		65.9 A	66.6 A	3.2 A	1.0 A	1.1 A
AYT-2_Exp4		67.3 B	67.7 B	3.2 A	1.0 A	1.1 A
AYT-2_Exp5		67.3 B	67.7 B	3.1 A	1.0 A	1.1 A
EYT_Exp4		67.5 B	68.2 B	3.1 A	1.0 A	1.0 A
EYT_Exp5		67.7 B	68.0 B	3.1 A	1.0 A	1.1 A

Levels not connected by the same letter are significantly different (Tukey α=0.01)

Relative efficacy among the five variety development experiments

The purpose of a variety development project is to identify the best lines through stages of field trials, where the expectation for each stage is to retain the best lines, while discarding lines that will not meet the breeding objectives. Because the best lines are among all lines in the initial stage, it is of interest to determine whether the best lines are retained or discarded after passing through all stages of development. None of the experiments retained all 10 of the best lines across all stages of development (Table 35). In experiment 1, from among the 10 best lines that entered the PYT, at most only 2, on average, were retained for all combinations of genetic

architectures and heritabilities. In contrast, experiments 2, 3, and 4 there was evidence of a slight improvement in retention of the best lines, especially for simulated yields with greater heritabilities on an entry mean basis. The ability to retain the best lines in experiment five was not as good as experiments 2, 3, and 4 except for the genetic architecture with a uniform distribution of simulated genetic effects.

Table 35: Average number of lines selected in EYT for each genetic architecture if the best 10 highest yield at PYT would have been selected for each experiment.

Experiment	3-Categories QTL			Uniform			Negative Exponential		
	Heritability								
	0.2	0.6	1.0	0.2	0.6	1.0	0.2	0.6	1.0
Exp-1	0	0	2	0	1	1	2	0	1
Exp-2	1	2	7	1	2	7	1	2	7
Exp-3	1	2	8	0	2	8	0	1	8
Exp-4	1	3	9	0	2	8	1	2	8
Exp-5	0	1	4	2	3	3	1	2	3

Relative efficiencies between variety development experiments 1 and 2

In order for experiment 1 and 2 to have the same costs, genotyping and selection using markers that are tightly linked to alleles responsible for SCN and PRR needs to be no more expensive than field based phenotyping for these same traits (Table 36).

Table 36: Resources allocated by task and generations are represented by US dollars and number of plots for experiments 1 and 2. MAS is practiced in F_2 generation for experiment 2.

Generation	Task	Families, lines or plants	Experiment 1		Experiment 2	
			$	Plots	$	Plots
F_1	selfing	300	2,250	75	2,250	75
F_2	selfing	300	9,000	300	9,000	300
F_2	genotyping	52,200			26,100	870
F_3	selfing	300	9,000	300	9,000	300
F_4	selfing	300	9,000	300	9,000	300
PYT	field trial	15,375	115,313	3,844	115,313	3,844
AYT-1	field trial	1,540	231,000	7,700	231,000	7,700
AYT-2	field trial	300	216,000	7,200	216,000	7,200
AYT-2	phenotyping*	300	26,100	870		
EYT	field trial	40	72,000	2,400	72,000	2,400
Total budget			**$689,663**	**22,989**	**$689,663**	**22,989**

* SCN and PRR greenhouse phenotyping

Relative efficiencies between variety development experiments 2 and 3

Experiments 2 and 3 have the same costs associated with the same tasks resulting in the same number of plots. However, experiment 3 has an added cost associated with genotyping 1,000 markers in the 112 founder lines. The costs associated with phenotyping the founder lines are applied to a previous cycle of variety development. And there are costs associated with conducting the data analyses, that are usually hidden in the overhead costs associated with plot costs (Table 37).

Table 37: Resources allocated by task and generations are represented by US dollars and number of plots for experiments 2 and 3. MAS is practiced in F_2 generation.

Generation	Task	Families, lines or plants	Experiment 2		Experiment 3	
			$	Plots	$	Plots
Founder	genotyping	112			2,240	75
F_1	selfing	300	2,250	75	2,250	75
F_2	selfing	300	9,000	300	9,000	300
F_2	genotyping	52,200	26,100	870	26,100	870
F_3	selfing	300	9,000	300	9,000	300
F_4	selfing	300	9,000	300	9,000	300
PYT	field trial	15,375	115,313	3,844	113,072	3,769
AYT-1	field trial	1,540	231,000	7,700	231,000	7,700
AYT-2	field trial	300	216,000	7,200	216,000	7,200
EYT	field trial	40	72,000	2,400	72,000	2,400
Total budget			**$689,663**	**22,989**	**$689,662**	**22,989**

Relative efficiencies between variety development experiments 3 and 4

Even though experiments 3 and 4 are initiated with different sets of founders, both are based on training sets involving the same numbers of founders and both require the same generations and tasks to achieve the releases from EYT, resulting in the same cost and number of plots (Table 38).

Table 38: Resources allocated by task and generations are represented by US dollars and number of plots for experiments 3 and 4.

Generation	Task	Families, lines or plants	Experiment 3		Experiment 4	
			$	Plots	$	Plots
Founder	genotyping	112	2,240	75	2,240	75
F_1	selfing	300	2,250	75	2,250	75
F_2	selfing	300	9,000	300	9,000	300
F_2	genotyping	52,200	26,100	870	26,100	870
F_3	selfing	300	9,000	300	9,000	300
F_4	selfing	300	9,000	300	9,000	300
PYT	field trial	15,375	113,072	3,769	113,072	3,769
AYT-1	field trial	1,540	231,000	7,700	231,000	7,700
AYT-2	field trial	300	216,000	7,200	216,000	7,200
EYT	field trial	40	72,000	2,400	72,000	2,400
Total budget			**$689,662**	**22,989**	**$689,662**	**22,989**

Relative efficiencies between variety development experiments 4 and 5

Experiment 5 has an added cost associated with genotyping 1,000 markers among 1,500 lines in a training set from an earlier generation of field trials from the previous generation in order to calculate GEBVs for PYTs. There is also an added cost of genotyping all lines that would have otherwise been grown in the PYT. The costs associated with phenotyping the training set are applied to a previous cycle of variety development. Break-even in this simulation was obtained with genotyping cost at $4.65, which allows keeping the same number of lines being tested from PYT to EYT stages at a total budget around $690k and 23k plots (Table 39).

Table 39: Resources allocated by task are represented by US dollars and number of plots for experiments 4 and 5.

Generation	Task	Families, lines or plants	Experiment 4		Experiment 5	
			$	Plots	$	Plots
Training Set	genotyping	1,500			30,000	1,000
F_1	selfing	300	2,250	75	2,250	75
F_2	selfing	300	9,000	300	9,000	300
F_2	genotyping	52,200	$26,100	870	$26,100	870
F_3	selfing	300	9,000	300	9,000	300
F_4	selfing	300	9,000	300	9,000	300
PYT	genotyping (GS)	15,375			71,490	2,383
PYT	seed increase	1,848			13,860	462
PYT	field trial	15,375	115,313	3,844		
AYT-1	field trial	1,540	231,000	7,700	231,000	7,700
AYT-2	field trial	300	216,000	7,200	216,000	7,200
EYT	field trial	40	72,000	2,400	72,000	2,400
Total budget			**$689,663**	**22,989**	**$689,700**	**22,990**

Relative efficiencies between variety development experiments 1 and 5

In order to understand the cumulative impacts of all marker assisted breeding strategies, we compared experiments 1 and 5 (Table 40). Experiment 5 has several added costs associated with genotyping, although the cost of conducting the PYT is reduced to seed increase in the same generation. The costs associated with phenotyping the training sets are applied to a previous cycle of variety development. Break-even was obtained at $4.65 per sample, which resulted in retention of the same number of lines being from PYT to EYT stages at a total budget around $690k and 23k plots.

Table 40: Resources allocated by task are represented by US dollars and number of plots for experiments 1 and 5. GS is practiced in PYT for experiment 5 at $4.65/genotyping.

Generation	Task	Families, lines or plants	Experiment 1		Experiment 5	
			$	Plots	$	Plots
Training Set	genotyping	1,540			30,000	1,000
F_1	selfing	300	2,250	75	2,250	75
F_2	selfing	300	9,000	300	9,000	300
F_2	genotyping	52,200			$26,100	870
F_3	selfing	300	9,000	300	9,000	300
F_4	selfing	300	9,000	300	9,000	300
PYT	genotyping (GS)	15,375			71,490	2,383
PYT	seed increase	1,848			13,860	462
PYT	field trial	15,375	115,313	3,844		
AYT-1	field trial	1,540	231,000	7,700	231,000	7,700
AYT-2	field trial	300	216,000	7,200	216,000	7,200
AYT-2	phenotyping*	300	26,100	870		
EYT	field trial	40	72,000	2,400	72,000	2,400
Total budget			**$689,663**	**22,989**	**$689,700**	**22,990**

* SCN and PRR greenhouse phenotyping

DISCUSSION

Variety development project assumptions

In recent decades soybean variety development projects have adopted emerging technologies such as computational logistics to assure timely delivery of seed from off-season nurseries to decrease time required to develop new varieties and marker assisted selection strategies to increase effectiveness of selection. During the same time genetic simulation models have been proposed as an important tool for evaluation of novel selection strategies (Podlich and Cooper, 1999; Heffner et al, 2010; Sun et al, 2011; Longin et al, 2015). To our knowledge, however, marker based technologies were adopted and integrated into variety development projects primarily as a reaction to perceived activities by competitors (Ted Crosbie, personal communication). The five simulated variety development experiments described herein were conducted as a post-hoc analysis of the impact of several marker enabled selection strategies into soybean variety development projects. Based on opinions of experienced soybean breeders these technologies began to be integrated into early stages of variety development projects as they existed at the beginning of this century. Because soybean variety development projects based on phenotypic selection at the beginning of this century were based on historical strategic decisions unique to every program, there are nuanced differences among projects. However, because all soybean variety development projects have the same constraints imposed by reproductive biology, our simulation model, based on a generic variety development pipeline (Figure 4), likely produced outcomes (as measured by genetic potential, sensitivity, specificity and costs) that are transferable to most soybean variety development projects. At a minimum, we hope that our results will encourage all plant breeders to utilize simulations with appropriate decision metrics

prior to making strategic decisions about integrating new methods and technologies into existing breeding programs.

Experiment 1 was designed to represent soybean variety development projects that existed at the turn of this century. Because it relies exclusively on phenotypic data for selection, it can be referred to as a conventional variety development project and provided us with a 'baseline' from which to compare other innovations in marker enhanced selection strategies. It deviates from actual projects at the turn of the century in two significant ways. First, it does not include introgression of transgenes such as herbicide resistance. Second, we assumed that phenotypes for SCN and PRR were expressed at the appropriate stages (AYT-2) through use of special nurseries and/or greenhouse experiments.

A commonly held opinion about MAS for oligogenic traits prior to the PYT is that selection will reduce the genetic potential for yield of lines evaluated in subsequent field trials and thus reduce yields of varieties emerging from the pipeline. On the other hand, there is an opinion that waiting until AYT-2 to select for secondary traits such as resistance to SCN and PRR will result in high yielding lines with no resistance to these pests. Our results did not support either opinion, although we did not investigate the impact of sample size on the outcomes. Translating these opinions into testable hypotheses needs to be conditional on the number of segregating lines or number of field plots. Future simulation research could determine the minimal sample sizes and numbers of field plots per stage to assure that MAS prior to the PYT will not affect yields of varieties released from the project.

The implementation of selecting crosses based on genotypic information parameters for experiment 3 intended to maximize the probability of producing better varieties. Mohammadi et al. (2015) stated that the ability of predictions has to discriminate among crosses with similar

predicted high means but different levels of genetic variance. According to Bernardo (2010), desirable breeding populations have both a high mean and a large genetic variance. However, this becomes more complex when multiple traits are assumed to be important in a crop, and some traits are negatively or uncorrelated at all.

Results in this study did not find any significant difference for outcomes between experiment 2 and 3. However, the outcomes from the EYT's of experiments 3 and 4 were significantly different, demonstrating that use of a training set composed of lines developed from a previous cycle of variety development is much better than using a set of random subset of desirable crosses. We hypothesize that if the training set is expanded to include genotypic and phenotypic data from all lines in the AYT-1 and AYT-2, not merely those involved in the EYT stage will further improve the ability to select crosses for initiating a variety development project.

Because heritabilities of GEBV's are less than heritabilities of phenotypic BV's (Dekkers, 2007), we decided to investigate whether utilizing GEBV's for selection during the PYT stage would adversely affect outcomes from the EYT stage. Experiments 4 and 5 were designed to compare phenotypic selection and genomic selection in PYTs. For all nine combinations of genetic architecture and heritability, the prediction accuracy estimated from cross-validation decreased substantially for heritabilities of 20%, and 60% on an entry mean basis. This is consistent with other studies e.g., Moser et al. (2015). However, results of genotypic values for yield from the EYT stage indicated similar genetic potential from both GEBV and phenotypic selection systems as well as the effectiveness of decision metrics.

Other publications have also shown the viability of using GEBVs in plant breeding programs (Poland et al. 2012; Bao et al. 2014; Jarquín et al. 2014; Spindel et al. 2015). Crossa et

al. (2014) mentioned how genomic selection have been explored at CIMMYT's wheat and maize breeding programs, and the emphasis given to genomic information improvement, especially on pedigree structure and validation set. Increasing training sets is one way to improve prediction accuracies (Gaynor et al. 2017). He et al. (2016) observed prediction accuracy on yield of 0.47 corresponding to field trials conducted at 3–4 locations, and suggested GS for yield is an interesting alternative for replacing early stages of selection in soybean breeding.

In soybean breeding, decreasing generation interval is a challenge because reproductive biology technologies such as double-haploids are not yet fully available. Therefore, application of GS to speed up multiple cycles of genetic improvement in a breeding program may have significant impacts not observed in a variety development project associated with materials from a single cycle of genetic improvement. Crossa et al. (2014) suggested that the greatest potential for GS is to use it if traditional phenotypic systems are expensive, time consuming or not biologically or logistically possible. Ceron-Rojas et al. (2015) observed phenotypic selection to be more accurate and with higher selection response per selection cycle than GS, but in terms of genetic gain per unit of time GS was better. Asoro et al. (2013) found a small advantage of MAS and GS over phenotypic selection on a per cycle basis, but the more pronounced benefit of GS is the selection response from conducting two cycles per year, which is very limited or even impossible for phenotypic selection in temperate environments.

A related question is how early can GS be implemented? In experiment 5, GEBVs were estimated for PYT lines ($F_{4:5}$), which brings an opportunity to grow these lines in off-season locations. A more aggressive use of GS was proposed by McClosky et al. (2013) and Hickey et al. (2014) in which they suggested using GEBVs for selection of individual plants in early segregating generations. Gaynor et al. (2017) suggest using GS for population improvement

through rapid recurrent selection and to identify new varieties or parents using traditional breeding designs. They found that applying GS resulted in reducing cycle time from 4 to 3 years per cycle of genetic improvement compared to a conventional system without GS, even though GS was less accurate than phenotypic selection.

A new simulation experiment is proposed to investigate the impact of applying MAS and GS in PYT eliminating MAS in F_2. Similarly, we recommend simulations with unbalanced data, which is more realistic with a real soybean project, which normally leads to lower repeatability, with direct impact on selection.

Genetic architecture

Genetic architectures for simply inherited and oligogenic traits are beginning to be well understood and thus can be modeled with some confidence in simulations of breeding projects. However, the genetic architectures for polygenic traits such as yield are not understood with any level of confidence. Estimates of identified QTLs for polygenic traits such as yield seem to be dependent on genetic background (Zhang et al. 2017) and environmental influences (Wang et al. 2014). In soybeans, 188 QTLs have been identified for yield (Sebastian et al. 2010; Fox et al. 2015). Also, more than 300 QTLs have been identified for seed weight, a yield component (https://www.soybase.org/). Spindel et al. (2015) suggest that genetic architecture for yield is best modeled using an infinitesimal model. Even though non-additive gene action is assumed negligible in self-pollinated crops (Bernardo 2003), a 5% increment on prediction accuracy was observed when shifting models from additive effects to additive plus epistatic effects (He et al. 2016). Herein we considered three models for yield, each consisting primarily of small additive effects distributed randomly throughout the genome. Even though we used three distinctive genetic architectures composed of additive genetic effects for yield the outcomes were similar.

Due to the limited recombination that occurs within a breeding cycle (variety development stages all occur within a single breeding cycle), it is likely that we could have approximated the infinitesimal model with only 300 QTL evenly distributed among the 40 independently segregating chromosome arms. Also with this finite number of haplotypes, there are likely non-linear contributions from epistasis, and genotype *x* environmental interactions. We leave these genetic architecture issues for future research on optimization of variety development projects.

Effectiveness metrics

The traditional way to report prediction accuracy ($r_{\ddot{y}g}$) as Pearson's correlation between the phenotype and GEBV ($r_{\ddot{y}y}$) divided by *h* (*h* is the square root of H^2) is not meaningful when breeders have to select or discard lines among multiple stages of field trials. For instance, one project with multiple traits can be represented by 1,000 lines at one stage of development. The breeder can only select 10% of those 1,000 lines to replicate in another year of trials. Our results show that the specificity metric is a more appropriate metric for decision makers.

Overall, sensitivity results were smaller than specificity in all experiments which is in agreement with decades of soybean breeding showing the challenge of selecting superior lines. On the other hand, specificity and decision accuracy results were high across all stages of selection, regardless of the experiment or the underlying genetic architecture and heritability of the trait. This is an artifact of calculations in which a very small proportion of lines are retained. Nonetheless, specificity results provided an opportunity to generate hypotheses about saving resources, because a researcher is able to discard undesirable lines before evaluating lines in more expensive later trial stages. Moreover, decision accuracy presented medium-high values, but this was mostly affected due to specificity than sensitivity. According to He et al. (2016), early selection stages is characterized by negative selection, i.e., discarding low performance

lines, while in more advanced stages is practiced positive selection for superior lines. Among the three metrics, specificity is more meaningful in PYT and AYT-1 stages where there are large number of lines with undesirable trait values.

Further research is needed to determine the optimal selection intensities at each stage of development. Such research should utilize Receiver Operator Curve (ROC) and Area Under Curve (AUC) analyses to determine the optimal selection intensities in the different testing stages. A ROC is based on "separator" scale, which provides a pair of overlapping distributions (Hajian-Tilaki, 2013). The plot is showed at several thresholds (i.e., selection intensity, number of markers, population size) and then the AUC is calculated as the accuracy of the prediction (Yang et al. 2016), using AUC analysis. Furthermore, the most desirable property of ROC and AUC analyses is that the optimal thresholds for selection intensities can be determined using objective criteria (Hajian-Tilaki, 2013). It may be that 50 years of trial and error have enabled soybean breeders to arrive at optimal thresholds, but this conjecture needs to be evaluated objectively.

Cost effectiveness

Previously, Xu and Crouch (2008) and Heslot et al (2015) have recognized trade-offs associated with costs among phenotypic and markers based selection strategies. The trade-off between application of MAS in F_2 (experiment 2 to 5) versus phenotypic screening in AYT-2 (experiment 1) demonstrated that MAS has an advantage of only advancing individuals or lines across stages of development if they carry desirable traits, but the F_2 family size can be a challenge if a great number of traits are under selection. On the other hand, experiment 1 does not identify lines with undesirable trait values until AYT-2 stage, wasting resources in testing and operational aspects such as seed production of lines that will be discarded. However, further

investigation is suggested to understand the different scenarios where single or more complex traits are incorporated to a breeding project. Assuming the break-even calculations are reasonably correct, are the current F_2 population size for experiments 2 to 5 (174 plants in this project) sufficient for selecting individuals carrying new favorable alleles? In terms of resources, what is the impact of increasing numbers of F_2's per cross?

Our results indicated several benefits from using GS in soybean which can be expanded with more research in this area and development of computational techniques or new markers technologies. We show that experiment 5 can be implemented at the same cost of experiment 1, without make any change to the project if genotyping cost decrease to a value of about US$4.65 per sample. On the other hand, if genotyping cost is maintained at the current level or increases, then different strategies will be required to incorporate GS in a breeding project.

More expensive genotyping technologies will directly affect the number of lines in testing stages. For example, the current cost of $20 per sample for GBS and DaRT based marker systems will force breeders to decrease AYT-1 and AYT-2 evaluations by about 50%, to achieve the break-even of a total budget around $690k and 23k plots. If larger selection intensities were applied in the PYT stage, then the numbers of plots needed for the AYT stages can be reduced. However, what would be the impact on discarding the best lines and reducing useful genetic variability in subsequent cycles? These are questions that we encourage researchers to investigate in future studies supporting GS be effectively applied to breeding and variety development.

REFERENCES

Alliprandini, L. F., Abatti, C., Bertagnolli, P. F., Cavassim, J. E., Gabe, H. L., Kurek, A., Matsumoto, M.N., de Oliveira, M.A.R., Pitol, C., Prado, L.C. and Steckling, C. 2009. Understanding soybean maturity groups in Brazil: environment, cultivar classification, and stability. *Crop science*, *49*(3), 801-808.

Anand, S. C., and Torrie, J. H. 1963. Heritability of yield and other traits and interrelationships among traits in the F3 and F4 generations of three soybean crosses. *Crop Science*, *3*(6), 508-511.

Arelli, P. R., Pantalone, V. R., Allen, F. L., Mengistu, A., & Fritz, L. A. 2015. Registration of JTN-5203 soybean germplasm with resistance to multiple cyst nematode populations. *Journal of Plant Registrations*, *9*(1), 108-114.

Asoro, F. G., Newell, M. A., Beavis, W. D., Scott, M. P., Tinker, N. A., & Jannink, J. L. 2013. Genomic, marker-assisted, and pedigree-BLUP selection methods for β-glucan concentration in elite oat. *Crop Science*, *53*(5), 1894-1906.

Beavis, W. D. 1994. The power and deceit of QTL experiments: lessons from comparative QTL studies. In *Proceedings of the forty-ninth annual corn and sorghum industry research conference* Chicago: American Seed Trade Association; 1994:250–266

Bao, Y., Vuong, T., Meinhardt, C., Tiffin, P., Denny, R., Chen, S., ... & Young, N. D. 2014. Potential of association mapping and genomic selection to explore PI 88788 derived soybean cyst nematode resistance. *The Plant Genome*, *7*(3).

Bernardo, R. 2003. Parental selection, number of breeding populations, and size of each population in inbred development. *Theoretical and applied genetics*, *107*(7), 1252-1256.

Bernardo, R., 2010. *Breeding for quantitative traits in plants* (p. 390). Woodbury: Stemma press.

Bernardo, R., 2014. Genomewide selection when major genes are known. *Crop Science*, *54*(1), pp.68-75.

Bonnett, D. G., Rebetzke, G. J., and Spielmeyer, W. 2005. Strategies for efficient implementation of molecular markers in wheat breeding. *Molecular Breeding*, *15*(1), 75-85.

Cardinal, A.J., Wang, S., Bowman, D.T. and Pantalone, V.R., 2012. Registration of 'NC-Burton'Soybean. *Journal of Plant Registrations*, *6*(2), pp.146-149.

Carter, A.H., Kidwell, K.K., DeMacon, V., Balow, K.A., Shelton, G.B., Higginbotham, R.W., Chen, X.M., Engle, D.A., Baik, B. and Morris, C.F., 2015. Registration of 'Sprinter'hard red winter wheat. *Journal of Plant Registrations*, *9*(2), pp.196-200.

Ceron-Rojas, J. J., Crossa, J., Arief, V. N., Basford, K., Rutkoski, J., Jarquín, D., ... & DeLacy, I. 2015. A genomic selection index applied to simulated and real data. *G3: Genes, Genomes, Genetics*, *5*(10), 2155-2164.

Cianzio, S. R., Lundeen, P., Gebhart, G., Rivera-Velez, N., Bhattacharyya, M. K., and Swaminathan, S. 2016. Registration of AR11SDS Soybean Germplasm Resistant to Sudden Death Syndrome, Soybean Cyst Nematode, and with Moderate Iron Deficiency Chlorosis Scores. *Journal of Plant Registrations*, *10*(2), 177-188.

Collard, B. C. Y., Jahufer, M. Z. Z., Brouwer, J. B., & Pang, E. C. K. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica*, *142*(1-2), 169-196.

Crossa, J., Pérez, P., Hickey, J., Burgueño, J., Ornella, L., Cerón-Rojas, J., ... & Bonnett, D. 2014. Genomic prediction in CIMMYT maize and wheat breeding programs. *Heredity*, *112*(1), 48-60.

Dekkers, J. C. M. 2007. Prediction of response to marker-assisted and genomic selection using selection index theory. *Journal of Animal Breeding and Genetics*, *124*(6), 331-341.

Diers, B. W., Kim, K. S., Frederick, R. D., Hartman, G. L., Unfried, J., Schultz, S., and Cary, T. 2014. Registration of eight soybean germplasm lines resistant to soybean rust. *Journal of Plant Registrations*, *8*(1), 96-101.

Endelman, J. B. 2011. Ridge regression and other kernels for genomic selection with R package rrBLUP. *The Plant Genome*, 4(3), 250-255.

Falconer, D. S., & Mackay, T. F. 1996. C. 1996. *Introduction to quantitative genetics*, *4*.

FAOSTAT. 2017. FAO statistics database. http://www.fao.org/faostat/en/#compare. (accessed 07 July 2017).

Fox, C. M., Cary, T. R., Nelson, R. L., & Diers, B. W. 2015. Confirmation of a seed yield QTL in soybean. *Crop Science*, *55*(3), 992-998.

Fraser, A., & Burnell, D. 1970. Computer models in genetics. *Computer models in genetics.*

Gaynor, R. C., Gorjanc, G., Bentley, A. R., Ober, E. S., Howell, P., Jackson, R., ... & Hickey, J. M. 2017. A two-part strategy for using genomic selection to develop inbred lines. *Crop Science*.

Goodman, M., & Lasker, G. W. 1974. Measurement of distance and propinquity in anthropological studies. In *Genetic Distance* (pp. 5-21). Springer US.

Grant, D., Nelson, R. T., Cannon, S. B., & Shoemaker, R. C. 2009. SoyBase, the USDA-ARS soybean genetics and genomics database. *Nucleic acids research*, *38*(suppl_1), D843-D846. http://soybase.org/. (accessed 05 March 2007).

Guo, B., Wang, D., Guo, Z., and Beavis, W. D. 2013. Family-based association mapping in crop species. *Theoretical and applied genetics*, *126*(6), 1419-1430.

Hajian-Tilaki, K. 2013. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian journal of internal medicine*, *4*(2), 627.

Hazel, L. N. 1943. The genetic basis for constructing selection indexes. *Genetics*, *28*(6), 476-490.

He, S., Schulthess, A. W., Mirdita, V., Zhao, Y., Korzun, V., Bothe, R., ... & Jiang, Y. 2016. Genomic selection in a commercial winter wheat population. *Theoretical and applied genetics*, *129*(3), 641-651.

Heffner, E. L., Lorenz, A. J., Jannink, J. L., & Sorrells, M. E. 2010. Plant breeding with genomic selection: gain per unit time and cost. *Crop science*, *50*(5), 1681-1690.

Heslot, N., Jannink, J. L., & Sorrells, M. E. 2015. Perspectives for genomic selection applications and research in plants. *Crop Science*, *55*(1), 1-12.

Hickey, J. M., Chiurugwi, T., Mackay, I., & Powell, W. 2017. Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. *Nature genetics*, *49*(9), 1297.

Howard, R., Carriquiry, A. L., & Beavis, W. D. 2014. Parametric and nonparametric statistical methods for genomic selection of traits with additive and epistatic genetic architectures. *G3: Genes, Genomes, Genetics*, *4*(6), 1027-1046.

Howard, R., Carriquiry, A. L., & Beavis, W. D. 2017. Application of response surface ethods to determine conditions for optimal genomic prediction. *G3: Genes, Genomes, Genetics*, *7*(9), 3103-3113.

Jarquín, D., Kocak, K., Posadas, L., Hyma, K., Jedlicka, J., Graef, G., and Lorenz, A. 2014. Genotyping by sequencing for genomic prediction in a soybean breeding population. *BMC genomics*, *15*(1), 740.

Johnson, H. W., Robinson, H. F., and Comstock, R. E. 1955. Genotypic and phenotypic correlations in soybeans and their implications in selection. *Agronomy journal*, *47*(10), 477-483.

Li, X., Zhu, C., Wang, J. and Yu, J., 2012. Computer simulation in plant breeding. *Annual Review of Genetics*, *35*, pp.303-339.

Lian, L., Jacobson, A., Zhong, S., and Bernardo, R. 2014. Genomewide prediction accuracy within 969 maize biparental populations. *Crop Science*, *54*(4), 1514-1522.

Longin, C. F. H., Mi, X., & Würschum, T. 2015. Genomic selection in wheat: optimum allocation of test resources and comparison of breeding strategies for line and hybrid breeding. *Theoretical and applied genetics*, *128*(7), 1297-1306.

Meuwissen, T. H., Hayes, B. J., & Goddard, M. E. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819.

McClosky, B., LaCombe, J., & Tanksley, S. D. 2013. Selfing for the design of genomic selection experiments in biparental plant populations. *Theoretical and applied genetics*, *126*(11), 2907-2920.

Mikel M.A., Diers B.W., Nelson R.L. and Smith H.H. 2010. Genetic diversity and agronomic improvement of North American soybean germplasm. *Crop Science* 50: 1219-1229.

Minella, E., Eichelberger, L., Costamilan, L.M. and Scagliusi, S.M., 2015. BRS Sampa: malting barley cultivar for irrigated production. *Crop Breeding and Applied Biotechnology*, *15*(1), pp.43-44.

Mohammadi, M., Tiede, T., & Smith, K. P. 2015. PopVar: a genome-wide procedure for predicting genetic variance and correlated response in biparental breeding populations. *Crop Science*, 55(5), 2068-2077.

Morrison, M. J., Voldeng, H. D., and Cober, E. R. 2000. Agronomic changes from 58 years of genetic improvement of short-season soybean cultivars in Canada. *Agronomy Journal*, *92*(4), 780-784.

Moser, G., Lee, S. H., Hayes, B. J., Goddard, M. E., Wray, N. R., and Visscher, P. M. 2015. Simultaneous discovery, estimation and prediction analysis of complex traits using a Bayesian mixture model. *PLoS genetics*, *11*(4), e1004969.

Neus, J. 2010. Marker assisted selection for seed yield in soybean [*Glycine max* (L.) Merr.] plant row yield trials (Dissertation, Graduate College of the University of Illinois at Urbana-Champaign). Retrieved from https://www.ideals.illinois.edu/bitstream/handle/2142/24173/Neus_Jason.pdf?sequence=1

Nei, M., & Li, W. H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences*, *76*(10), 5269-5273.

Niblack, T. L. 2005. Soybean cyst nematode management reconsidered. *Plant disease* 89.10: 1020-1026.

Orf, J. H., Chase, K., Jarvik, T., Mansur, L. M., Cregan, P. B., Adler, F. R., and Lark, K. G. 1999. Genetics of soybean agronomic traits: I. Comparison of three related recombinant inbred populations. *Crop Science*, *39*(6), 1642-1651.

Pengyin, C. H. E. N. 2017. *U.S. Patent No. 9,622,444*. Washington, DC: U.S. Patent and Trademark Office.

Podlich, D. W., & Cooper, M. 1999. Modelling plant breeding programs as search strategies on a complex response surface. In *Asia-Pacific Conference on Simulated Evolution and Learning* (pp. 171-178). Springer, Berlin, Heidelberg.

Poland, J., Endelman, J., Dawson, J., Rutkoski, J., Wu, S., Manes, Y., Dreisigacker, S., Crossa, J., Sánchez-Villeda, H., Sorrells, M., and Jannink, J.L. 2012. Genomic Selection in Wheat Breeding using Genotyping-by-Sequencing. The Plant Genome Journal, 5(3): 103.

Recker, J. R., Burton, J. W., Cardinal, A., and Miranda, L. 2014. Genetic and phenotypic correlations of quantitative traits in two long-term, randomly mated soybean populations. *Crop Science*, *54*(3), 939-943.

Riedelsheimer, C., Endelman, J. B., Stange, M., Sorrells, M. E., Jannink, J. L., & Melchinger, A. E. 2013. Genomic predictability of interconnected biparental maize populations. *Genetics*, *194*(2), 493-503.

Sebastian, S. A., Streit, L. G., Stephens, P. A., Thompson, J. A., Hedges, B. R., Fabrizius, M. A., ... and Feng, L. 2010. Context-specific marker-assisted selection for improved grain yield in elite soybean populations. *Crop science*, *50*(4), 1196-1206.

Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., ... & Xu, D. 2009. Genome sequence of the palaeopolyploid soybean. *Nature (London)*, *463*(LBNL-6794E).

Shultz, J. L., Kurunam, D., Shopinski, K., Iqbal, M. J., Kazi, S., Zobrist, K., ... & Yesudas, C. R. 2006. The Soybean Genome Database (SoyGD): a browser for display of duplicated, polyploid, regions and sequence tagged sites on the integrated physical and genetic maps of Glycine max. *Nucleic Acids Research*, *34*(suppl_1), D758-D765.

Simko, I., & Pechenick, D. A. 2010. Combining partially ranked data in plant breeding and biology: I. Rank aggregating methods. *Communications in Biometry & Crop Science*, *5*(1).

Smith, H. F. 1936. A discriminant function for plant selection. *Annals of Human Genetics*, *7*(3), 240-250.

Song, Q., Jenkins, J., Jia, G., Hyten, D. L., Pantalone, V., Jackson, S. A., and Cregan, P. B. 2016. Construction of high resolution genetic linkage maps to improve the soybean genome sequence assembly Glyma1. 01. *BMC genomics*, 17(1), 33.

Specht, J.E., Diers, B.W., Nelson, R.L., Ferraz de Toledo, J.F., Torrion, J.A. and Grassini, P. 2014. CSSA Special Publications, Yield Gains in Major U.S. Field Crops, 33:311-356.

Spindel, J., Begum, H., Akdemir, D., Virk, P., Collard, B., Redoña, E., ... & McCouch, S. R. 2015. Genomic selection and association mapping in rice (Oryza sativa): effect of trait genetic architecture, training population composition, marker number and statistical model on accuracy of rice genomic selection in elite, tropical rice breeding lines. *PLoS genetics*, *11*(2), e1004982.

Sun, X., Peng, T. and Mumm, R.H. 2011. The role and basics of computer simulation in support of critical decisions in plant breeding. *Molecular breeding*, *28*(4), pp.421-436.

Thompson, J. A., and Nelson, R. L. 1998. Utilization of diverse germplasm for soybean yield improvement. *Crop science*, *38*(5), 1362-1368.

USDA-FAS. 2016. Oilseeds: World Markets and Trade. https://apps.fas.usda.gov/psdonline/circulars/oilseeds.pdf (accessed 25 July 2017).

Walling, J. G., Shoemaker, R., Young, N., Mudge, J., & Jackson, S. 2006. Chromosome-level homeology in paleopolyploid soybean (Glycine max) revealed through integration of genetic and chromosome maps. *Genetics*, *172*(3), 1893-1900.

Wang, X., Pang, Y., Zhang, J., Zhang, Q., Tao, Y., Feng, B., ... & Li, Z. 2014. Genetic background effects on QTL and QTL× environment interaction for yield and its component traits as revealed by reciprocal introgression lines in rice. *The Crop Journal*, *2*(6), 345-357.

Wang, J., Van Ginkel, M., Podlich, D., Ye, G., Trethowan, R., Pfeiffer, W., DeLacy, I.H., Cooper, M. and Rajaram, S. 2003. Comparison of two breeding strategies by computer simulation. *Crop Science*, *43*(5), pp.1764-1773.

Werner, K., Friedt, W., and Ordon, F. 2005. Strategies for pyramiding resistance genes against the barley yellow mosaic virus complex (BaMMV, BaYMV, BaYMV-2). *Molecular Breeding*, *16*(1), 45-55.

Würschum, T., Reif, J. C., Kraft, T., Janssen, G., & Zhao, Y. 2013. Genomic selection in sugar beet breeding populations. *BMC genetics*, *14*(1), 85.

Xu, Y., & Crouch, J. H. 2008. Marker-assisted selection in plant breeding: from publications to practice. *Crop science*, *48*(2), 391-407.

Yang, S., Fresnedo-Ramírez, J., Wang, M., Cote, L., Schweitzer, P., Barba, P., ... & Sacks, G. 2016. A next-generation marker genotyping platform (AmpSeq) in heterozygous crops: a case study for marker-assisted selection in grapevine. *Horticulture research*, *3*, 16002.

Zhang, H., Song, Q., Griffin, J. D., & Song, B. H. 2017. Genetic architecture of wild soybean (Glycine soja) response to soybean cyst nematode (Heterodera glycines). *Molecular Genetics and Genomics*, 1-9.

Zhong, S., & Jannink, J. L. 2007. Using quantitative trait loci results to discriminate among crosses on the basis of their progeny mean and variance. *Genetics*, *177*(1), 567-576.

APPENDIX A1

LIST OF 112 PUBLIC VARIETIES, INCLUDING 27 KEY GERMPLASM AND 85 COMMERCIAL VARIETIES GROWN FROM 1995 TO 2008.

Access Number	Access Name	Date of acquisition	PI information
PI548573	Harosoy	1956	Key germplasm
PI548506	Amsoy	1962	Key germplasm
PI548628	Wayne	1962	Key germplasm
PI548540	Corsoy	1963	Key germplasm
PI548660	Bragg	1963	Key germplasm
PI548551	Hark	1965	Key germplasm
PI548527	Calland	1966	Key germplasm
PI548510	Beeson	1967	Key germplasm
PI548547	Cutler	1967	Key germplasm
PI553039	Davis	1967	Key germplasm
PI548631	Williams	1970	Key germplasm
PI548655	Forrest	1972	Key germplasm
PI548667	Essex	1972	Key germplasm
PI548561	Hodgson	1973	Key germplasm
PI548679	Mitchell	1973	Key germplasm
PI556511	A3127	1977	Key germplasm
PI548523	Pella	1979	Key germplasm
PI518674	Fayette	1980	Key germplasm
PI518664	Hutcheson	1988	Key germplasm
PI539936	A4715	1990	Key germplasm
PI548689	B216	1990	Key germplasm
PI548974	Bedford	1991	Key germplasm
PI548975	Centennial	1991	Key germplasm
PI548983	Tracy	1991	Key germplasm
PI548989	Ransom	1991	Key germplasm
PI559370	Mack	1991	Key germplasm
PI556689	A2943	2003	Key germplasm
PI587185	Probst	1995	cultivar
PI593654	Stressland	1995	cultivar
PI593257	LN90-4524	1996	cultivar
PI593258	Macon	1996	cultivar
PI593259	Iroquois	1996	cultivar
PI595753	ODell	1996	cultivar
PI595754	Nemaha	1996	cultivar
PI595843	Flint	1996	cultivar
PI595926	Athow	1996	cultivar
PI586980	KS3494	1997	cultivar
PI586981	KS4694	1997	cultivar
PI593463	General	1997	cultivar
PI595081	KS4895	1997	cultivar
PI596407	Defiance	1997	cultivar

APPENDIX A1: CONTINUED

PI596412	LN92-12033	1997	cultivar
PI596413	LN92-12054	1997	cultivar
PI597381	Savoy	1997	cultivar
PI597382	Omaha	1997	cultivar
PI597383	LN89-3264	1997	cultivar
PI597384	LN89-3615	1997	cultivar
PI597386	Dwight	1997	cultivar
PI597387	Pana	1997	cultivar
PI598222	TN 4-94	1997	cultivar
PI599811	C1943	1997	cultivar
PI602996	LG91-7350R	1997	cultivar
PI602059	Apollo	1998	cultivar
PI602060	Olympus	1998	cultivar
PI602449	BARC-15 nodulated	1998	cultivar
PI602450	BARC-15 non-nodulated	1998	cultivar
PI604464	HC95-15MB	1998	cultivar
PI606748	REND	1998	cultivar
PI606749	INA	1998	cultivar
PI607385	LN92-7369	1999	cultivar
PI610670	NE3297	1999	cultivar
PI610671	NE3399	1999	cultivar
PI611112	7499	1999	cultivar
PI612594	Kottman	2000	cultivar
PI612931	HF93-083	2000	cultivar
PI612932	HF93-035	2000	cultivar
PI614088	Loda	2000	cultivar
PI614154	Darby	2000	cultivar
PI614155	HS93-4118	2000	cultivar
PI614806	Troll	2000	cultivar
PI614807	Stout	2000	cultivar
PI614808	Strong	2000	cultivar
PI614832	NE3400	2000	cultivar
PI615553	LG92-1255	2001	cultivar
PI615554	LG93-7054	2001	cultivar
PI615555	LG93-7654	2001	cultivar
PI615556	LG93-7792	2001	cultivar
PI620883	LS93-0375	2001	cultivar
PI632401	APEX	2002	cultivar
PI632402	STALWART	2002	cultivar
PI632422	HC94-1946	2002	cultivar
PI632424	HC94-35PR	2002	cultivar
PI632425	HC98-303	2002	cultivar
PI632426	HC96-45PR	2002	cultivar
PI632427	HC97-4358	2002	cultivar
PI632428	HC94-944	2002	cultivar

APPENDIX A1: CONTINUED

PI632429	HC95-933	2002	cultivar
PI632430	HC95-261PR	2002	cultivar
PI632431	HC94-168	2002	cultivar
PI556687	A3966	2003	cultivar
PI556729	A2522	2003	cultivar
PI633608	Dilworth	2003	cultivar
PI633729	LG92-4208	2003	cultivar
PI633730	LG94-1128	2003	cultivar
PI633731	LG94-1906	2003	cultivar
PI633732	LG94-4667	2003	cultivar
PI633983	LN97-15076	2003	cultivar
PI634335	LS94-3207	2003	cultivar
PI634827	NE2701	2004	cultivar
PI636695	S99-11509	2005	cultivar
PI639282	LG96-1797	2005	cultivar
PI639283	LG97-7012	2005	cultivar
PI639284	LG98-1445	2005	cultivar
PI639285	LG98-1605	2005	cultivar
PI639740	LD00-3309	2005	cultivar
PI640911	AxN-1-55	2005	cultivar
PI642055	DT97-4290	2006	cultivar
PI642768	OHIO FG5	2006	cultivar
PI644024	Stout-Rps1k	2006	cultivar
PI644025	Strong-Rps1k	2006	cultivar
PI643146	Prohio	2008	cultivar
PI654356	S99-2281	2008	cultivar

APPENDIX A2

SOYBEAN CHROMOSOMES AND LINKAGE GROUPS (LG), WITH THEIR RESPECTIVE SIZE IN CM AND NUMBER OF LOCI CHROMOSOME AND NUMBER OF YIELD LOCI PER CHROMOSOME.

Chromosome	**Chr length (cM)**	**# Loci/CHR (34,708)**	**Yield Loci/CHR (3,000)**	**Sample (2,301loci)**	**LG**
Gm01	98.4	1,354	114	99	D1a
Gm02	140.6	2,032	182	141	D1b
Gm03	99.5	1,449	129	100	N
Gm04	112.3	1,667	120	113	C1
Gm05	86.8	1,505	127	87	A1
Gm06	136.5	1,602	143	137	C2
Gm07	135.2	1,788	156	136	M
Gm08	146.7	2,112	186	147	A2
Gm09	99.6	1,665	140	100	K
Gm10	132.9	1,709	152	133	O
Gm11	124.2	1,374	126	125	B1
Gm12	120.5	1,238	111	121	H
Gm13	120.0	2,284	200	120	F
Gm14	108.2	1,584	151	109	B2
Gm15	99.9	2,099	189	100	E
Gm16	92.3	1,522	112	93	J
Gm17	119.2	1,580	128	120	D2
Gm18	105.0	2,916	260	105	G
Gm19	101.1	1,929	163	102	L
Gm20	112.8	1,299	111	113	I
TOTAL	2291.6	34,708	3,000	2,301	

APPENDIX A3

ANALYSIS OF VARIANCE FOR SIMULATED GENOTYPIC VALUES OF YIELD DUE TO FOUNDER SETS, STAGE OF DEVELOPMENT, GENETIC ARCHITECTURES AND HERITABILITY IN EXPERIMENT 1: PHENOTYPIC SELECTION FOR ALL TRAITS.

Source	**Sum of Squares**	**Mean Square**	**F Ratio**	**Prob > F**
Model	19414.6	242.6	237.3	<.0001*
Error	9578.3	1.0		
C. Total	28992.9			
Source		**S.Squares**	**F Ratio**	**Prob > F**
Founder_Set		14.6	7.1	0.0008*
Stage		480.9	235.2	<.0001*
Stage*Founder_Set		2.1	0.5	0.7117
Genetic_Architecture		1365.5	667.8	<.0001*
Stage*Genetic_Architecture		210.2	51.4	<.0001*
Founder_Set*Genetic_Architecture		10.3	2.5	0.0379*
Stage*Founder_Set*Genetic_Architecture		3.7	0.4	0.8839
Heritability		3040.6	1487.0	<.0001*
Stage*Heritability		26.1	6.4	<.0001*
Founder_Set*Heritability		22.4	5.4	0.0002*
Stage*Founder_Set*Heritability		1.9	0.2	0.9829
Genetic_Architecture*Heritability		388.9	95.1	<.0001*
Stage*Genetic_Architecture*Heritability		25.3	3.0	0.0017*
Founder_Set*Genetic_Architecture*Heritability		16.3	2.0	0.0421*
Stage*Founder_Set*Genetic_Archit* Heritability		2.4	0.1	1.0000

APPENDIX A4

SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT OF SET 1 FROM EXPERIMENT 1.

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield				Lodging				SCN				PRR			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	**0.2**	64.3	64.3	64.1	63.8	66.6	66.3	66.2	66.0	2.9	2.7	3.1	3.2	5.3	3.6	4.6	3.1	3.3	3.5	3.8	3.7
		0.6	63.7	64.5	65.2	64.8	64.9	65.7	66.4	66.0	3.3	3.3	3.6	3.5	4.4	3.5	2.8	2.6	3.4	3.7	3.6	4.1
		1.0	64.1	66.1	66.2	67.2	64.1	66.1	66.2	67.2	2.9	2.9	3.1	2.9	4.7	5.1	4.6	3.3	3.9	3.3	3.8	4.0
	uniform	**0.2**	62.0	62.0	62.1	61.9	63.5	63.3	63.2	63.0	2.9	3.1	2.8	2.9	4.7	4.7	4.1	4.2	4.3	4.5	4.1	3.5
		0.6	63.4	63.4	63.4	63.2	64.0	63.9	63.8	63.7	3.3	3.3	3.4	3.2	6.3	6.8	7.2	5.7	3.1	3.0	3.1	3.3
		1.0	65.3	65.2	65.0	64.7	65.3	65.2	65.0	64.7	3.2	2.9	3.1	3.4	4.7	4.8	5.2	4.6	3.1	3.7	4.1	3.6
	Negative Exponential	**0.2**	61.6	61.8	61.6	61.7	62.7	62.7	62.7	62.6	3.2	3.1	3.1	3.1	5.6	4.7	6.4	5.4	3.3	3.7	3.6	3.8
		0.6	63.3	63.4	63.4	63.0	63.7	63.9	64.0	63.6	2.8	3.2	3.1	3.1	5.4	5.7	5.5	4.4	3.7	3.5	3.7	3.5
		1.0	65.5	66.2	66.3	66.2	65.5	66.2	66.3	66.2	3.1	3.3	3.2	3.3	5.7	5.3	5.8	6.0	3.2	4.1	3.5	4.3
AYT-2	3-categories QTL	**0.2**	65.2	65.4	64.9	64.3	65.8	66.2	65.2	65.2	2.6	2.4	3.3	3.2	1.0	1.0	1.0	1.0	1.4	1.8	1.2	1.6
		0.6	64.5	65.6	66.6	66.1	64.8	66.0	66.8	66.8	3.2	3.2	3.9	3.6	1.0	1.0	1.0	1.0	2.0	1.4	1.6	1.8
		1.0	65.0	66.4	67.3	69.1	65.0	66.4	67.3	69.1	2.8	3.1	3.6	3.5	1.0	1.0	1.0	1.0	1.4	1.4	1.2	2.0
	uniform	**0.2**	62.1	62.1	62.4	62.4	62.5	62.5	63.1	62.8	3.3	3.0	2.9	3.6	1.0	1.0	1.0	1.0	1.6	1.8	1.8	1.4
		0.6	63.2	63.4	63.2	63.5	63.1	63.8	63.3	63.6	3.4	4.0	3.9	2.7	1.0	1.0	1.0	1.0	1.8	1.4	1.8	1.2
		1.0	65.8	65.2	64.9	65.4	65.8	65.2	64.9	65.4	3.6	2.5	3.6	3.7	1.0	1.0	1.0	1.0	1.2	1.4	1.8	2.0
	Negative Exponential	**0.2**	61.8	62.2	61.8	61.8	62.2	62.7	62.0	62.4	3.5	3.2	3.2	2.7	1.0	1.0	1.0	1.0	1.4	1.4	1.4	1.8
		0.6	63.8	63.5	64.0	63.4	63.8	63.7	63.8	63.6	2.9	3.5	3.4	3.5	1.0	1.0	1.0	1.0	1.8	2.0	1.6	2.4
		1.0	66.0	66.6	66.4	66.3	66.0	66.6	66.4	66.3	3.1	3.9	3.6	3.2	1.0	1.0	1.0	1.0	1.4	2.2	1.6	1.2
EYT	3-categories QTL	**0.2**	65.6	65.7	65.7	64.3	66.5	67.2	66.4	66.5	2.3	2.0	4.0	3.0	1.0	1.0	1.0	1.0	1.0	2.0	1.0	2.3
		0.6	65.4	66.3	67.4	66.7	65.8	66.6	68.6	67.6	4.0	4.0	3.3	4.0	1.0	1.0	1.0	1.0	3.0	1.0	1.7	1.0
		1.0	66.3	67.2	68.7	70.3	66.3	67.2	68.7	70.3	2.0	3.0	4.3	3.7	1.0	1.0	1.0	1.0	3.0	1.7	1.0	1.7
	uniform	**0.2**	62.7	62.9	62.5	62.8	62.7	63.3	63.2	63.5	5.0	4.0	3.0	3.7	1.0	1.0	1.0	1.0	3.0	1.7	2.3	1.0
		0.6	63.7	63.9	63.9	63.9	64.1	64.0	64.4	64.3	3.7	4.3	5.0	3.0	1.0	1.0	1.0	1.0	1.7	1.0	2.0	1.0
		1.0	66.4	65.7	65.6	65.9	66.4	65.7	65.6	65.9	3.7	1.3	3.0	5.0	1.0	1.0	1.0	1.0	1.0	1.0	3.0	1.7
	Negative Exponential	**0.2**	62.1	62.3	61.7	62.2	63.1	63.0	62.5	63.2	3.3	3.7	2.5	4.0	1.0	1.0	1.0	1.0	1.7	1.7	2.0	2.0
		0.6	64.1	63.9	64.6	64.0	64.4	64.6	64.7	64.5	3.5	3.7	4.3	2.5	1.0	1.0	1.0	1.0	2.0	2.3	1.7	3.0
		1.0	67.3	67.7	67.2	67.2	67.3	67.7	67.2	67.2	2.5	4.3	3.7	3.0	1.0	1.0	1.0	1.0	1.0	1.7	1.7	2.0

APPENDIX A5

SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT OF SET 2 FROM EXPERIMENT 1.

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield				Lodging				SCN				PRR			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	**0.2**	64.0	64.2	64.0	63.9	66.2	66.1	66.4	66.1	2.8	2.8	2.9	3.0	3.8	4.2	4.1	4.1	3.9	3.6	3.7	4.1
		0.6	63.8	64.5	64.8	64.8	64.9	65.7	66.0	66.0	3.3	3.5	3.3	3.5	4.1	2.9	3.0	3.0	3.2	3.9	3.5	3.8
		1.0	63.4	64.7	65.9	67.2	63.4	64.7	65.9	67.2	2.9	3.2	3.0	3.0	4.5	4.2	3.6	3.0	3.8	4.0	3.5	3.8
	uniform	**0.2**	62.1	62.2	62.0	62.0	63.3	63.4	63.2	63.3	2.9	3.1	2.9	3.1	4.9	4.2	3.3	3.5	4.6	4.5	4.1	3.1
		0.6	63.6	63.6	63.8	63.6	64.1	64.1	64.3	64.0	3.0	3.0	3.0	3.6	6.2	7.2	7.3	6.9	3.9	4.1	3.3	3.3
		1.0	65.6	65.0	65.1	64.9	65.6	65.0	65.1	64.9	3.2	3.1	3.5	3.3	5.6	5.8	5.4	4.1	3.6	3.1	4.3	4.0
	Negative Exponential	**0.2**	61.6	61.7	61.8	61.8	62.6	62.6	62.8	62.8	3.1	3.5	3.1	3.1	6.2	5.8	5.8	6.1	4.1	4.0	4.0	4.0
		0.6	63.3	63.2	63.2	63.0	63.9	63.7	63.8	63.6	2.7	3.2	3.2	3.0	5.7	5.7	4.9	5.6	4.0	3.6	4.3	3.6
		1.0	65.2	66.0	66.1	66.0	65.2	66.0	66.1	66.0	3.0	3.6	3.2	3.3	5.3	4.9	5.7	5.8	3.7	4.7	4.2	3.9
AYT-2	3-categories QTL	**0.2**	64.7	65.5	64.7	64.3	65.1	66.4	65.5	64.9	2.2	2.6	3.4	2.9	1.0	1.0	1.0	1.0	1.6	1.2	1.2	1.4
		0.6	64.7	65.9	66.2	65.8	65.2	66.2	66.6	66.6	3.2	3.8	3.3	3.5	1.0	1.0	1.0	1.0	1.6	1.2	1.8	1.4
		1.0	64.9	65.7	66.5	68.9	64.9	65.7	66.5	68.9	2.9	3.3	3.1	3.2	1.0	1.0	1.0	1.0	1.8	2.0	1.8	1.6
	uniform	**0.2**	62.0	62.2	62.5	62.3	62.2	62.5	63.1	62.8	3.4	3.0	3.4	3.4	1.0	1.0	1.0	1.0	2.2	1.6	2.0	1.6
		0.6	63.4	63.6	64.0	63.4	63.3	63.8	64.0	63.3	3.5	3.2	3.3	3.1	1.0	1.0	1.0	1.0	3.0	2.2	2.2	1.4
		1.0	65.7	65.3	65.0	65.6	65.7	65.3	65.0	65.6	3.3	2.8	3.4	3.1	1.0	1.0	1.0	1.0	1.4	1.8	2.2	1.6
	Negative Exponential	**0.2**	61.6	61.9	62.0	61.8	61.9	62.1	62.0	62.2	2.8	3.2	2.7	3.1	1.0	1.0	1.0	1.0	1.0	1.6	1.2	1.8
		0.6	63.2	63.6	63.6	63.1	63.2	63.8	63.6	63.4	3.2	3.6	3.4	3.2	1.0	1.0	1.0	1.0	1.4	2.0	2.0	1.6
		1.0	65.6	66.2	66.1	66.0	65.6	66.2	66.1	66.0	3.2	3.6	3.6	3.5	1.0	1.0	1.0	1.0	1.8	2.8	2.2	2.4
EYT	3-categories QTL	**0.2**	65.6	66.1	65.0	65.0	65.8	66.9	66.9	66.8	2.0	3.7	3.0	2.0	1.0	1.0	1.0	1.0	1.7	1.7	1.0	1.0
		0.6	65.7	66.8	67.5	65.4	66.7	67.4	68.1	66.7	3.7	4.0	3.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.7	3.0
		1.0	67.3	66.8	67.2	69.5	67.3	66.8	67.2	69.5	2.0	3.5	3.0	3.0	1.0	1.0	1.0	1.0	2.0	2.0	1.0	2.3
	uniform	**0.2**	62.6	63.4	62.5	62.8	63.1	62.9	63.6	63.1	3.5	3.0	3.0	2.7	1.0	1.0	1.0	1.0	1.5	1.0	1.0	1.0
		0.6	63.7	64.0	64.7	63.5	64.2	64.5	65.0	63.9	4.7	2.7	3.3	3.0	1.0	1.0	1.0	1.0	1.7	1.0	3.0	1.0
		1.0	66.4	65.9	65.6	66.1	66.4	65.9	65.6	66.1	4.0	2.0	3.0	3.3	1.0	1.0	1.0	1.0	1.7	1.7	1.0	1.0
	Negative Exponential	**0.2**	61.9	62.0	62.4	62.1	63.9	62.5	63.4	63.2	3.0	4.0	4.0	3.3	1.0	1.0	1.0	1.0	1.0	1.7	1.0	1.0
		0.6	63.6	63.9	64.0	64.3	64.2	64.2	64.5	63.8	3.3	3.7	4.0	4.0	1.0	1.0	1.0	1.0	1.0	2.3	1.7	1.0
		1.0	67.1	66.8	66.8	66.8	67.1	66.8	66.8	66.8	3.0	3.3	3.3	2.3	1.0	1.0	1.0	1.0	1.0	2.3	1.7	2.3

APPENDIX A6

SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT OF SET 3 FROM EXPERIMENT 1.

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield				Lodging				SCN				PRR			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	64.4	64.4	64.3	63.8	66.3	66.6	66.4	65.9	2.9	2.7	2.9	3.1	4.3	3.9	4.2	4.1	3.6	3.3	3.6	4.1
		0.6	63.5	64.2	65.2	65.2	64.7	65.3	66.2	66.2	3.3	3.4	3.2	3.4	3.3	3.1	3.5	2.7	3.4	3.8	3.4	3.6
		1.0	62.4	64.6	66.0	66.6	62.4	64.6	66.0	66.6	2.9	2.9	2.9	3.1	3.8	3.5	3.7	4.5	3.2	4.5	4.1	4.3
	uniform	0.2	62.0	62.1	62.1	62.2	63.2	63.4	63.4	63.3	3.1	2.9	3.0	2.6	5.4	3.7	4.2	5.1	3.7	4.1	3.9	2.8
		0.6	63.1	63.6	63.4	63.3	63.6	64.0	63.9	63.8	3.2	3.0	3.1	3.4	6.4	7.4	6.0	6.7	4.4	3.6	3.5	4.3
		1.0	65.3	65.0	64.9	64.7	65.3	65.0	64.9	64.7	3.1	3.4	3.2	3.0	4.4	4.4	5.4	4.0	4.1	4.7	4.8	3.3
	Negative Exponential	0.2	61.6	61.6	61.7	61.7	62.7	62.8	62.9	62.7	3.0	3.2	3.1	3.3	6.1	4.8	5.4	5.2	3.9	4.2	3.5	3.3
		0.6	62.9	63.2	63.4	62.8	63.3	63.7	63.8	63.4	3.1	3.0	3.1	3.2	4.3	4.7	4.8	5.4	3.9	4.1	3.7	3.9
		1.0	64.9	65.8	66.2	66.0	64.9	65.8	66.2	66.0	3.1	3.2	3.0	3.2	4.5	5.3	6.1	6.3	3.7	4.5	3.9	4.2
AYT-2	3-categories QTL	0.2	65.1	64.9	65.1	63.9	65.7	65.8	65.4	64.7	2.5	2.7	3.2	2.4	1.0	1.0	1.0	1.0	1.6	1.6	1.4	1.2
		0.6	64.5	65.4	66.5	66.5	65.0	65.4	66.8	67.2	3.4	3.6	3.1	3.3	1.0	1.0	1.0	1.0	1.6	1.0	1.4	1.6
		1.0	63.6	65.1	66.4	67.4	63.6	65.1	66.4	67.4	3.1	3.3	3.1	3.3	1.0	1.0	1.0	1.0	1.2	1.8	1.8	1.2
	uniform	0.2	62.2	62.4	62.6	62.6	62.8	62.5	62.5	63.1	3.7	2.6	3.1	2.6	1.0	1.0	1.0	1.0	1.4	1.8	1.2	1.4
		0.6	63.0	63.6	63.5	63.3	63.1	63.8	63.6	63.2	3.1	2.8	3.0	3.6	1.0	1.0	1.0	1.0	2.4	1.4	1.4	1.4
		1.0	65.5	65.0	65.0	65.3	65.5	65.0	65.0	65.3	3.6	3.3	2.8	2.6	1.0	1.0	1.0	1.0	2.2	1.8	3.0	1.4
	Negative Exponential	0.2	61.4	61.9	62.0	61.8	61.6	62.2	62.3	62.2	3.0	3.5	3.2	3.2	1.0	1.0	1.0	1.0	1.6	1.4	2.0	1.6
		0.6	63.4	63.2	63.5	62.9	63.5	63.5	63.7	63.0	3.2	2.8	3.0	3.2	1.0	1.0	1.0	1.0	1.6	1.6	1.4	1.8
		1.0	65.3	65.6	66.2	65.9	65.3	65.6	66.2	65.9	3.4	3.2	3.4	3.1	1.0	1.0	1.0	1.0	1.6	1.6	1.6	2.4
EYT	3-categories QTL	0.2	65.5	65.3	65.6	64.9	66.0	66.3	66.3	67.8	2.7	3.0	3.7	2.0	1.0	1.0	1.0	1.0	1.7	1.7	1.0	1.0
		0.6	65.8	67.1	66.7	66.7	67.0	67.2	67.5	68.0	4.0	4.0	4.0	3.3	1.0	1.0	1.0	1.0	1.0	1.0	1.7	1.7
		1.0	65.3	66.2	67.5	69.3	65.3	66.2	67.5	69.3	5.0	3.7	3.0	3.3	1.0	1.0	1.0	1.0	1.0	1.7	1.0	1.0
	uniform	0.2	62.7	62.9	62.9	62.6	63.7	63.7	63.2	63.2	3.0	2.0	2.7	2.7	1.0	1.0	1.0	1.0	1.0	2.3	1.7	1.7
		0.6	63.9	64.0	63.8	63.5	64.0	64.4	64.1	63.6	3.5	2.3	3.3	4.0	1.0	1.0	1.0	1.0	3.0	1.0	1.0	1.0
		1.0	66.2	65.5	65.7	65.8	66.2	65.5	65.7	65.8	3.0	2.3	3.0	3.3	1.0	1.0	1.0	1.0	2.3	1.0	3.0	1.0
	Negative Exponential	0.2	62.0	62.1	62.1	61.9	62.9	62.8	62.6	62.6	1.5	3.0	4.0	2.7	1.0	1.0	1.0	1.0	2.0	1.0	1.7	1.7
		0.6	63.8	63.2	64.1	64.2	63.8	63.9	64.2	64.8	3.3	3.0	3.0	2.5	1.0	1.0	1.0	1.0	1.0	2.0	1.0	2.0
		1.0	66.5	66.9	67.0	66.6	66.5	66.9	67.0	66.6	4.5	4.5	3.3	3.3	1.0	1.0	1.0	1.0	2.0	1.0	1.7	3.0

APPENDIX A7

LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 1 (SET 1), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES.

Stage	Yield Gen. Architecture	H^2	Genotypic yield				Phenotypic yield			
			1.8-2.4	**2.4-3.0**	**3.0-3.6**	**3.6-4.2**	**1.8-2.4**	**2.4-3.0**	**3.0-3.6**	**3.6-4.2**
AYT-1	**3-categories QTL**	**0.2**	64.3 (±0.10)	64.3 (±0.11)	64.1 (±0.10)	63.8 (±0.11)	66.6 (±0.09)	66.3 (±0.10)	66.2 (±0.09)	66.0 (±0.10)
		0.6	63.7	64.5	65.2	64.8	64.9	65.7	66.4	66.0
		1.0	64.1	66.1	66.2	67.2	64.1	66.1	66.2	67.2
	uniform	**0.2**	62.0	62.0	62.1	61.9	63.5	63.3	63.2	63.0
		0.6	63.4	63.4	63.4	63.2	64.0	63.9	63.8	63.7
		1.0	65.3	65.2	65.0	64.7	65.3	65.2	65.0	64.7
	Negative Exponential	**0.2**	61.6	61.8	61.6	61.7	62.7	62.7	62.7	62.6
		0.6	63.3	63.4	63.4	63.0	63.7	63.9	64.0	63.6
		1.0	65.5	66.2	66.3	66.2	65.5	66.2	66.3	66.2
AYT-2	**3-categories QTL**	**0.2**	65.2 (±0.21)	65.4 (±0.20)	64.9 (±0.20)	64.3 (±0.21)	65.8 (±0.21)	66.2 (±0.18)	65.2 (±0.22)	65.2 (±0.20)
		0.6	64.5	65.6	66.6	66.1	64.8	66.0	66.8	66.8
		1.0	65.0	66.4	67.3	69.1	65.0	66.4	67.3	69.1
	uniform	**0.2**	62.1	62.1	62.4	62.4	62.5	62.5	63.1	62.8
		0.6	63.2	63.4	63.2	63.5	63.1	63.8	63.3	63.6
		1.0	65.8	65.2	64.9	65.4	65.8	65.2	64.9	65.4
	Negative Exponential	**0.2**	61.8	62.2	61.8	61.8	62.2	62.7	62.0	62.4
		0.6	63.8	63.5	64.0	63.4	63.8	63.7	63.8	63.6
		1.0	66.0	66.6	66.4	66.3	66.0	66.6	66.4	66.3
EYT	**3-categories QTL**	**0.2**	65.6 (±0.30)	65.7 (±0.27)	65.7 (±0.26)	64.3 (±0.43)	66.5 (±0.36)	67.2 (±0.26)	66.4 (±0.29)	66.5 (±0.41)
		0.6	65.4	66.3	67.4	66.7	65.8	66.6	68.6	67.6
		1.0	66.3	67.2	68.7	70.3	66.3	67.2	68.7	70.3
	uniform	**0.2**	62.7	62.9	62.5	62.8	62.7	63.3	63.2	63.5
		0.6	63.7	63.9	63.9	63.9	64.1	64.0	64.4	64.3
		1.0	66.4	65.7	65.6	65.9	66.4	65.7	65.6	65.9
	Negative Exponential	**0.2**	62.1	62.3	61.7	62.2	63.1	63.0	62.5	63.2
		0.6	64.1	63.9	64.6	64.0	64.4	64.6	64.7	64.5
		1.0	67.3	67.7	67.2	67.2	67.3	67.7	67.2	67.2

APPENDIX A8

LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 1 (SET 2), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES.

Stage	Yield Gen. Architecture	H^2	Genotypic yield				Phenotypic yield			
			1.8-2.4	**2.4-3.0**	**3.0-3.6**	**3.6-4.2**	**1.8-2.4**	**2.4-3.0**	**3.0-3.6**	**3.6-4.2**
AYT-1	**3-categories QTL**	**0.2**	64.0 (±0.11)	64.2 (±0.09)	64.0 (±0.10)	63.9 (±0.10)	66.2 (±0.09)	66.1 (±0.08)	66.4 (±0.09)	66.1 (±0.10)
		0.6	63.8	64.5	64.8	64.8	64.9	65.7	66.0	66.0
		1.0	63.4	64.7	65.9	67.2	63.4	64.7	65.9	67.2
	uniform	**0.2**	62.1	62.2	62.0	62.0	63.3	63.4	63.2	63.3
		0.6	63.6	63.6	63.8	63.6	64.1	64.1	64.3	64.0
		1.0	65.6	65.0	65.1	64.9	65.6	65.0	65.1	64.9
	Negative Exponential	**0.2**	61.6	61.7	61.8	61.8	62.6	62.6	62.8	62.8
		0.6	63.3	63.2	63.2	63.0	63.9	63.7	63.8	63.6
		1.0	65.2	66.0	66.1	66.0	65.2	66.0	66.1	66.0
AYT-2	**3-categories QTL**	**0.2**	64.7 (±0.25)	65.5 (±0.20)	64.7 (±0.20)	64.3 (±0.19)	65.1 (±0.26)	66.4 (±0.18)	65.5 (±0.20)	64.9 (±0.18)
		0.6	64.7	65.9	66.2	65.8	65.2	66.2	66.6	66.6
		1.0	64.9	65.7	66.5	68.9	64.9	65.7	66.5	68.9
	uniform	**0.2**	62.0	62.2	62.5	62.3	62.2	62.5	63.1	62.8
		0.6	63.4	63.6	64.0	63.4	63.3	63.8	64.0	63.3
		1.0	65.7	65.3	65.0	65.6	65.7	65.3	65.0	65.6
	Negative Exponential	**0.2**	61.6	61.9	62.0	61.8	61.9	62.1	62.0	62.2
		0.6	63.2	63.6	63.6	63.1	63.2	63.8	63.6	63.4
		1.0	65.6	66.2	66.1	66.0	65.6	66.2	66.1	66.0
EYT	**3-categories QTL**	**0.2**	65.6 (±0.26)	66.1 (±0.26)	65.0 (±0.30)	65.0 (±0.21)	65.8 (±0.26)	66.9 (±0.23)	66.9 (±0.35)	66.8 (±0.36)
		0.6	65.7	66.8	67.5	65.4	66.7	67.4	68.1	66.7
		1.0	67.3	66.8	67.2	69.5	67.3	66.8	67.2	69.5
	uniform	**0.2**	62.6	63.4	62.5	62.8	63.1	62.9	63.6	63.1
		0.6	63.7	64.0	64.7	63.5	64.2	64.5	65.0	63.9
		1.0	66.4	65.9	65.6	66.1	66.4	65.9	65.6	66.1
	Negative Exponential	**0.2**	61.9	62.0	62.4	62.1	63.9	62.5	63.4	63.2
		0.6	63.6	63.9	64.0	64.3	64.2	64.2	64.5	63.8
		1.0	67.1	66.8	66.8	66.8	67.1	66.8	66.8	66.8

APPENDIX A9

LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 1 (SET 3), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES.

Stage	Yield Gen. Architecture	H^2	Genotypic yield				Phenotypic yield			
			1.8-2.4	**2.4-3.0**	**3.0-3.6**	**3.6-4.2**	**1.8-2.4**	**2.4-3.0**	**3.0-3.6**	**3.6-4.2**
AYT-1	**3-categories QTL**	**0.2**	64.4 (±0.09)	64.4 (±0.10)	64.3 (±0.10)	63.8 (±0.10)	66.3 (±0.09)	66.6 (±0.09)	66.4 (±0.09)	65.9 (±0.09)
		0.6	63.5	64.2	65.2	65.2	64.7	65.3	66.2	66.2
		1.0	62.4	64.6	66.0	66.6	62.4	64.6	66.0	66.6
	uniform	**0.2**	62.0	62.1	62.1	62.2	63.2	63.4	63.4	63.3
		0.6	63.1	63.6	63.4	63.3	63.6	64.0	63.9	63.8
		1.0	65.3	65.0	64.9	64.7	65.3	65.0	64.9	64.7
	Negative Exponential	**0.2**	61.6	61.6	61.7	61.7	62.7	62.8	62.9	62.7
		0.6	62.9	63.2	63.4	62.8	63.3	63.7	63.8	63.4
		1.0	64.9	65.8	66.2	66.0	64.9	65.8	66.2	66.0
AYT-2	**3-categories QTL**	**0.2**	65.1 (±0.21)	64.9 (±0.23)	65.1 (±0.19)	63.9 (±0.23)	65.7 (±0.21)	65.8 (±0.20)	65.4 (±0.21)	64.7 (±0.23)
		0.6	64.5	65.4	66.5	66.5	65.0	65.4	66.8	67.2
		1.0	63.6	65.1	66.4	67.4	63.6	65.1	66.4	67.4
	uniform	**0.2**	62.2	62.4	62.6	62.6	62.8	62.5	62.5	63.1
		0.6	63.0	63.6	63.5	63.3	63.1	63.8	63.6	63.2
		1.0	65.5	65.0	65.0	65.3	65.5	65.0	65.0	65.3
	Negative Exponential	**0.2**	61.4	61.9	62.0	61.8	61.6	62.2	62.3	62.2
		0.6	63.4	63.2	63.5	62.9	63.5	63.5	63.7	63.0
		1.0	65.3	65.6	66.2	65.9	65.3	65.6	66.2	65.9
EYT	**3-categories QTL**	**0.2**	65.5 (±0.34)	65.3 (±0.32)	65.6 (±0.29)	64.9 (±0.24)	66.0 (±0.43)	66.3 (±0.41)	66.3 (±0.30)	67.8 (±0.33)
		0.6	65.8	67.1	66.7	66.7	67.0	67.2	67.5	68.0
		1.0	65.3	66.2	67.5	69.3	65.3	66.2	67.5	69.3
	uniform	**0.2**	62.7	62.9	62.9	62.6	63.7	63.7	63.2	63.2
		0.6	63.9	64.0	63.8	63.5	64.0	64.4	64.1	63.6
		1.0	66.2	65.5	65.7	65.8	66.2	65.5	65.7	65.8
	Negative Exponential	**0.2**	62.0	62.1	62.1	61.9	62.9	62.8	62.6	62.6
		0.6	63.8	63.2	64.1	64.2	63.8	63.9	64.2	64.8
		1.0	66.5	66.9	67.0	66.6	66.5	66.9	67.0	66.6

APPENDIX A10

DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY IN FOUNDER SET 1 USED IN EXPERIMENT 1 FOR NINE COMBINATIONS OF GENETIC ARCHITECTURE AND HERITABILITY FOR EACH OF FOUR STAGES IN SOYBEAN VARIETY DEVELOPMENT.

Stage	Yield genetic architecture	H^2	Accuracy	Sensitivity	Specificity
PYT	3-categories QTL	0.20	0.88	0.38	0.93
		0.60	0.93	0.63	0.96
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.88	0.38	0.93
		0.60	0.93	0.65	0.96
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.88	0.39	0.93
		0.60	0.93	0.64	0.96
		1.00	1.00	1.00	1.00
AYT-1	3-categories QTL	0.20	0.79	0.47	0.87
		0.60	0.86	0.63	0.91
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.81	0.52	0.88
		0.60	0.87	0.68	0.92
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.81	0.52	0.88
		0.60	0.86	0.65	0.92
		1.00	1.00	1.00	1.00
AYT-2	3-categories QTL	0.20	0.93	0.73	0.96
		0.60	0.94	0.78	0.97
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.91	0.68	0.95
		0.60	0.97	0.90	0.98
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.93	0.75	0.96
		0.60	0.95	0.80	0.97
		1.00	1.00	1.00	1.00
EYT	3-categories QTL	0.20	0.75	0.50	0.83
		0.60	0.75	0.50	0.83
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.75	0.50	0.83
		0.60	0.85	0.70	0.90
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.70	0.40	0.80
		0.60	0.80	0.60	0.87
		1.00	1.00	1.00	1.00

APPENDIX A11

DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY IN FOUNDER SET 2 USED IN EXPERIMENT 1 FOR NINE COMBINATIONS OF GENETIC ARCHITECTURE AND HERITABILITY FOR EACH OF FOUR STAGES IN SOYBEAN VARIETY DEVELOPMENT.

Stage	Yield genetic architecture	H^2	Accuracy	Sensitivity	Specificity
		0.20	0.88	0.38	0.93
	3-categories QTL	0.60	0.93	0.63	0.96
		1.00	1.00	1.00	1.00
		0.20	0.88	0.38	0.93
PYT	Uniform	0.60	0.94	0.68	0.96
		1.00	1.00	1.00	1.00
		0.20	0.88	0.40	0.93
	Negative Exponential	0.60	0.93	0.65	0.96
		1.00	1.00	1.00	1.00
		0.20	0.79	0.46	0.87
	3-categories QTL	0.60	0.85	0.62	0.91
		1.00	1.00	1.00	1.00
		0.20	0.83	0.56	0.89
AYT-1	Uniform	0.60	0.89	0.71	0.93
		1.00	1.00	1.00	1.00
		0.20	0.81	0.51	0.88
	Negative Exponential	0.60	0.85	0.63	0.91
		1.00	1.00	1.00	1.00
		0.20	0.93	0.73	0.96
	3-categories QTL	0.60	0.95	0.80	0.97
		1.00	1.00	1.00	1.00
		0.20	0.91	0.68	0.95
AYT-2	Uniform	0.60	0.97	0.88	0.98
		1.00	1.00	1.00	1.00
		0.20	0.98	0.93	0.99
	Negative Exponential	0.60	0.96	0.85	0.98
		1.00	1.00	1.00	1.00
		0.20	0.75	0.50	0.83
	3-categories QTL	0.60	0.80	0.60	0.87
		1.00	1.00	1.00	1.00
		0.20	0.75	0.50	0.83
EYT	Uniform	0.60	0.75	0.50	0.83
		1.00	1.00	1.00	1.00
		0.20	0.80	0.60	0.87
	Negative Exponential	0.60	0.75	0.50	0.83
		1.00	1.00	1.00	1.00

APPENDIX A12

DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY IN FOUNDER SET 3 USED IN EXPERIMENT 1 FOR NINE COMBINATIONS OF GENETIC ARCHITECTURE AND HERITABILITY FOR EACH OF FOUR STAGES IN SOYBEAN VARIETY DEVELOPMENT.

Stage	Yield genetic architecture	H^2	Accuracy	Sensitivity	Specificity
PYT	3-categories QTL	0.20	0.88	0.39	0.93
		0.60	0.92	0.62	0.96
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.88	0.39	0.93
		0.60	0.93	0.66	0.96
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.88	0.40	0.93
		0.60	0.93	0.63	0.96
		1.00	1.00	1.00	1.00
AYT-1	3-categories QTL	0.20	0.79	0.47	0.87
		0.60	0.84	0.59	0.90
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.82	0.54	0.89
		0.60	0.88	0.68	0.92
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.81	0.51	0.88
		0.60	0.86	0.64	0.91
		1.00	1.00	1.00	1.00
AYT-2	3-categories QTL	0.20	0.91	0.68	0.95
		0.60	0.95	0.83	0.97
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.92	0.70	0.95
		0.60	0.98	0.93	0.99
		1.00	0.99	0.98	1.00
	Negative Exponential	0.20	0.95	0.80	0.97
		0.60	0.93	0.73	0.96
		1.00	1.00	1.00	1.00
EYT	3-categories QTL	0.20	0.80	0.60	0.87
		0.60	0.80	0.60	0.87
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.70	0.40	0.80
		0.60	0.80	0.60	0.87
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.70	0.40	0.80
		0.60	0.85	0.70	0.90
		1.00	1.00	1.00	1.00

APPENDIX A13

ANALYSIS OF VARIANCE FOR SIMULATED GENOTYPIC VALUES FOR YIELD DUE TO STAGE OF DEVELOPMENT, FOUNDER SETS, GENETIC ARCHITECTURES AND HERITABILITY OF LINES FOR EXPERIMENT 2: MAS FOR SCN AND PRR TRAITS + PHENOTYPIC SELECTION FOR YIELD.

Source	**Sum of Squares**	**Mean Square**	**F Ratio**	**Prob > F**
Model	22662.1	283.2	279.5	<.0001*
Error	9494.2	1.0		
C. Total	32156.4			
Source		**S.Squares**	**F Ratio**	**Prob > F**
Founder_Set		7.7	3.8	0.0214*
Stage		2423.4	1195.7	<.0001*
Founder_Set*Stage		2.7	0.6	0.6025
Genetic_architecture		2519.0	1242.8	<.0001*
Founder_Set* Genetic_architecture		8.2	2.0	0.0852
Stage* Genetic_architecture		396.6	97.8	<.0001*
Founder_Set*Stage* Genetic_architecture		4.9	0.6	0.7742
Heritability		3635.7	1793.8	<.0001*
Founder_Set* Heritability		19.0	4.7	0.0009*
Stage* Heritability		214.6	52.9	<.0001*
Founder_Set*Stage* Heritability		3.6	0.4	0.8899
Genetic_architecture* Heritability		184.5	45.5	<.0001*
Founder_Set* Genetic_architecture* Heritability		14.0	1.7	0.0866
Stage* Genetic_architecture* Heritability		103.9	12.8	<.0001*
Founder_Set*Stage* Genetic_archit.* Heritability		2.5	0.1	0.9999

APPENDIX A14

SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FOR SET 1 FROM EXPERIMENT 2.

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield				Lodging				SCN				PRR			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	64.4	64.7	64.4	64.2	66.5	66.4	66.5	66.0	2.7	3.1	2.8	3.1	1.0	1.0	1.0	1.0	1.1	1.1	1.1	1.0
		0.6	63.9	64.5	65.0	65.0	65.0	65.6	66.2	66.0	3.2	3.1	3.3	3.3	1.0	1.0	1.0	1.0	1.0	1.1	1.1	1.1
		1.0	63.9	66.0	66.3	67.0	63.9	66.0	66.3	67.0	2.9	3.1	3.0	3.3	1.0	1.0	1.0	1.0	1.1	1.1	1.0	1.2
	uniform	0.2	62.2	62.1	62.0	62.0	63.5	63.4	63.1	63.2	3.1	3.1	3.1	2.8	1.0	1.0	1.0	1.0	1.1	1.1	1.2	1.1
		0.6	63.0	62.9	63.1	63.1	63.3	63.3	63.6	63.4	3.3	3.3	3.4	3.6	1.0	1.0	1.0	1.0	1.1	1.2	1.1	1.1
		1.0	65.0	64.7	64.8	64.7	65.0	64.7	64.8	64.7	3.5	3.3	3.4	3.1	1.0	1.0	1.0	1.0	1.1	1.1	1.1	1.1
	Negative Exponential	0.2	61.6	61.7	61.6	61.6	62.5	62.7	62.7	62.4	3.1	3.1	3.4	3.0	1.0	1.0	1.0	1.0	1.0	1.1	1.1	1.1
		0.6	63.1	63.2	63.3	62.8	63.8	63.8	63.9	63.4	2.9	3.2	3.1	3.1	1.0	1.0	1.0	1.0	1.0	1.1	1.1	1.1
		1.0	65.0	65.4	65.7	65.9	65.0	65.4	65.7	65.9	3.2	3.3	2.7	3.3	1.0	1.0	1.0	1.0	1.1	1.1	1.1	1.1
AYT-2	3-categories QTL	0.2	65.4	65.8	66.1	65.3	66.0	66.9	66.6	66.2	2.6	3.1	2.2	3.4	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	65.5	66.1	67.0	67.0	66.0	66.7	67.7	67.5	3.3	3.3	3.7	3.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.4
		1.0	67.0	69.1	68.8	69.8	67.0	69.1	68.8	69.8	3.0	2.9	2.7	3.4	1.0	1.0	1.0	1.0	1.2	1.0	1.0	1.2
	uniform	0.2	62.9	62.8	62.8	62.9	63.7	63.4	63.4	63.5	3.1	3.2	3.2	2.8	1.0	1.0	1.0	1.0	1.0	1.2	1.4	1.2
		0.6	63.7	63.7	64.0	64.1	64.0	63.9	64.2	64.4	3.2	3.6	3.5	4.2	1.0	1.0	1.0	1.0	1.2	1.2	1.0	1.2
		1.0	66.3	66.2	66.0	65.9	66.3	66.2	66.0	65.9	3.2	3.4	3.2	2.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	0.2	62.0	62.3	62.4	62.2	62.8	63.0	63.1	63.1	2.5	3.3	3.2	2.5	1.0	1.0	1.0	1.0	1.0	1.2	1.0	1.4
		0.6	64.5	64.2	64.2	63.8	64.9	64.5	64.6	64.1	2.8	3.1	3.2	2.6	1.0	1.0	1.0	1.0	1.0	1.2	1.0	1.0
		1.0	66.3	66.6	67.0	67.0	66.3	66.6	67.0	67.0	3.0	3.2	2.5	3.4	1.0	1.0	1.0	1.0	1.0	1.0	1.2	1.0
EYT	3-categories QTL	0.2	66.0	66.2	66.9	65.3	67.1	67.4	67.9	67.0	2.3	3.5	4.3	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	66.2	66.9	67.8	68.0	67.4	67.9	68.3	68.9	4.0	2.7	3.7	3.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	69.3	71.0	70.8	71.0	69.3	71.0	70.8	71.0	3.0	2.3	2.0	2.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	0.2	63.4	63.4	62.8	63.4	64.0	63.7	64.2	64.0	2.5	3.5	3.3	3.3	1.0	1.0	1.0	1.0	1.0	1.0	2.3	1.0
		0.6	64.1	64.0	64.6	64.3	64.6	64.4	64.5	64.7	3.5	3.7	4.0	4.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	66.6	66.7	66.8	67.1	66.6	66.7	66.8	67.1	2.7	2.3	4.0	2.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	0.2	62.4	62.3	62.4	62.5	63.1	63.2	63.7	63.1	1.5	4.0	4.3	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	65.0	64.4	64.4	65.0	65.5	65.1	65.0	65.0	3.3	3.0	4.0	2.0	1.0	1.0	1.0	1.0	1.0	1.7	1.0	1.0
		1.0	67.2	66.9	67.7	67.5	67.2	66.9	67.7	67.5	4.0	2.0	3.0	3.3	1.0	1.0	1.0	1.0	1.0	1.0	1.7	1.0

APPENDIX A15

SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FOR SET 2 FROM EXPERIMENT 2.

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield				Lodging				SCN				PRR			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	64.5	64.8	64.3	63.8	66.7	66.5	66.6	65.7	2.6	2.9	3.0	2.9	1.0	1.0	1.0	1.0	1.1	1.1	1.2	1.2
		0.6	63.9	64.5	65.4	65.4	65.0	65.5	66.5	66.5	3.2	3.0	3.3	3.4	1.0	1.0	1.0	1.0	1.0	1.1	1.1	1.2
		1.0	63.9	64.7	66.1	67.2	63.9	64.7	66.1	67.2	3.0	3.2	3.1	3.0	1.0	1.0	1.0	1.0	1.1	1.2	1.2	1.1
	uniform	0.2	62.0	62.1	62.1	62.1	63.4	63.4	63.3	63.2	3.0	2.8	2.9	2.9	1.0	1.0	1.0	1.0	1.1	1.2	1.2	1.1
		0.6	63.0	63.2	63.2	63.0	63.5	63.5	63.6	63.4	3.1	3.0	3.2	3.5	1.0	1.0	1.0	1.0	1.0	1.1	1.2	1.1
		1.0	65.2	64.7	65.0	64.8	65.2	64.7	65.0	64.8	3.1	3.2	3.4	3.4	1.0	1.0	1.0	1.0	1.1	1.1	1.1	1.0
	Negative Exponential	0.2	61.6	61.8	61.6	61.6	62.5	62.7	62.8	62.5	3.2	3.2	3.2	3.2	1.0	1.0	1.0	1.0	1.1	1.1	1.2	1.3
		0.6	63.3	63.0	63.1	62.6	63.8	63.7	63.6	63.1	3.0	2.8	3.2	3.2	1.0	1.0	1.0	1.0	1.1	1.2	1.1	1.1
		1.0	65.2	65.4	65.8	65.8	65.2	65.4	65.8	65.8	3.3	3.3	2.9	3.4	1.0	1.0	1.0	1.0	1.0	1.0	1.1	1.1
AYT-2	3-categories QTL	0.2	65.4	66.0	65.6	65.5	66.9	66.7	66.5	66.4	2.3	2.7	3.7	3.2	1.0	1.0	1.0	1.0	1.2	1.2	1.2	1.0
		0.6	65.4	66.7	67.0	67.5	66.1	67.3	67.4	67.9	3.5	3.6	3.2	3.6	1.0	1.0	1.0	1.0	1.2	1.2	1.0	1.6
		1.0	67.3	67.4	69.1	70.1	67.3	67.4	69.1	70.1	2.7	3.0	2.7	3.9	1.0	1.0	1.0	1.0	1.0	1.0	1.2	1.0
	uniform	0.2	62.9	62.8	62.9	62.8	63.6	63.6	63.6	63.4	3.3	2.6	3.1	2.7	1.0	1.0	1.0	1.0	1.0	1.0	1.4	1.0
		0.6	63.9	64.1	64.1	63.9	64.0	64.4	64.4	64.1	3.4	3.1	3.0	4.3	1.0	1.0	1.0	1.0	1.0	1.0	1.2	1.0
		1.0	66.5	65.7	66.1	65.9	66.5	65.7	66.1	65.9	3.8	4.0	3.7	3.8	1.0	1.0	1.0	1.0	1.0	1.0	1.2	1.0
	Negative Exponential	0.2	62.0	62.3	62.1	62.1	62.8	63.0	62.9	62.9	2.8	3.6	3.2	2.6	1.0	1.0	1.0	1.0	1.0	1.2	1.0	1.4
		0.6	64.5	63.9	64.4	63.4	65.0	64.3	64.8	63.7	2.6	2.8	3.6	3.2	1.0	1.0	1.0	1.0	1.0	1.2	1.2	1.0
		1.0	66.5	66.5	66.9	66.8	66.5	66.5	66.9	66.8	3.2	3.7	3.0	3.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
EYT	3-categories QTL	0.2	66.0	66.4	66.8	66.0	67.6	66.6	68.4	66.4	2.3	3.3	5.0	2.7	1.0	1.0	1.0	1.0	1.7	1.0	1.0	1.0
		0.6	66.2	67.9	68.7	68.0	67.4	68.5	69.7	68.1	4.0	3.3	4.0	2.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	70.1	69.4	71.5	71.2	70.1	69.4	71.5	71.2	3.0	3.0	2.7	3.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	0.2	63.4	62.9	63.0	63.7	64.4	63.8	64.0	64.0	3.0	3.0	2.5	2.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	64.6	64.5	64.5	64.4	65.0	64.5	65.0	64.6	3.5	3.5	2.3	4.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	66.8	66.3	66.5	66.3	66.8	66.3	66.5	66.3	4.3	3.5	3.3	5.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	0.2	62.4	62.6	62.3	62.1	63.0	63.1	63.4	62.9	1.0	4.7	3.7	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	3.0
		0.6	65.2	64.3	64.4	64.2	65.7	64.8	65.0	64.3	3.7	3.0	3.0	3.0	1.0	1.0	1.0	1.0	1.0	1.7	1.7	1.0
		1.0	67.1	67.0	67.5	67.5	67.1	67.0	67.5	67.5	4.0	2.5	3.0	3.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

APPENDIX A16

SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FOR SET 3 FROM EXPERIMENT 2.

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield				Lodging				SCN				PRR			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	64.6	64.7	64.2	63.9	66.7	66.7	66.2	66.0	2.5	2.7	3.0	2.9	1.0	1.0	1.0	1.0	1.2	1.2	1.1	1.0
		0.6	64.2	64.6	65.3	65.3	65.2	65.5	66.5	66.4	3.1	3.2	3.2	3.2	1.0	1.0	1.0	1.0	1.0	1.2	1.1	1.1
		1.0	63.2	64.6	65.9	67.0	63.2	64.6	65.9	67.0	3.3	3.1	3.1	3.2	1.0	1.0	1.0	1.0	1.0	1.1	1.2	1.1
	uniform	0.2	62.1	62.1	62.2	62.3	63.3	63.4	63.4	63.6	2.9	3.0	3.0	2.8	1.0	1.0	1.0	1.0	1.2	1.3	1.1	1.2
		0.6	63.0	63.0	63.1	63.0	63.3	63.4	63.5	63.3	3.4	3.0	3.0	3.5	1.0	1.0	1.0	1.0	1.1	1.0	1.1	1.2
		1.0	65.2	64.5	64.5	64.6	65.2	64.5	64.5	64.6	3.3	3.1	3.3	3.2	1.0	1.0	1.0	1.0	1.0	1.1	1.1	1.1
	Negative Exponential	0.2	61.6	61.7	61.7	61.7	62.5	62.7	62.7	62.6	2.8	2.9	3.2	2.9	1.0	1.0	1.0	1.0	1.0	1.1	1.1	1.1
		0.6	63.2	63.1	63.0	62.8	63.7	63.6	63.7	63.4	3.1	3.1	3.1	3.1	1.0	1.0	1.0	1.0	1.1	1.2	1.1	1.1
		1.0	64.8	65.3	65.7	66.1	64.8	65.3	65.7	66.1	3.2	3.3	3.1	3.2	1.0	1.0	1.0	1.0	1.1	1.1	1.3	1.1
AYT-2	3-categories QTL	0.2	65.7	65.8	65.4	65.3	66.6	66.9	66.2	65.9	2.4	3.0	2.7	3.1	1.0	1.0	1.0	1.0	1.6	1.2	1.4	1.0
		0.6	65.8	66.0	67.4	67.5	66.6	66.7	68.2	67.9	3.8	3.3	3.8	2.6	1.0	1.0	1.0	1.0	1.0	1.0	1.2	1.0
		1.0	66.1	67.1	68.1	70.2	66.1	67.1	68.1	70.2	3.6	2.8	3.9	4.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	0.2	62.9	62.9	62.9	63.1	63.5	63.8	63.8	63.9	3.1	2.6	3.0	2.4	1.0	1.0	1.0	1.0	1.2	1.2	1.0	1.2
		0.6	63.8	64.0	63.9	63.8	64.1	64.3	64.3	64.1	3.0	3.2	3.0	3.4	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.2
		1.0	66.2	65.3	65.6	65.5	66.2	65.3	65.6	65.5	4.3	3.3	3.7	3.0	1.0	1.0	1.0	1.0	1.0	1.2	1.0	1.0
	Negative Exponential	0.2	62.0	62.2	62.4	62.3	62.9	62.9	63.1	63.1	2.5	2.7	3.2	2.8	1.0	1.0	1.0	1.0	1.0	1.2	1.2	1.4
		0.6	64.4	64.1	64.1	63.8	64.9	64.5	64.5	64.2	3.2	2.8	2.9	3.0	1.0	1.0	1.0	1.0	1.0	1.2	1.0	1.0
		1.0	65.9	66.6	66.8	67.5	65.9	66.6	66.8	67.5	3.0	3.4	3.1	3.0	1.0	1.0	1.0	1.0	1.2	1.0	1.4	1.0
EYT	3-categories QTL	0.2	65.9	66.0	65.9	65.9	66.8	66.8	67.3	66.1	3.3	2.3	2.0	3.0	1.0	1.0	1.0	1.0	3.0	1.7	1.0	1.0
		0.6	66.6	67.1	67.6	68.5	67.6	67.5	68.2	69.5	3.0	4.0	4.0	2.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	68.2	67.8	69.0	71.7	68.2	67.8	69.0	71.7	4.0	4.0	3.0	3.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	0.2	63.4	63.1	63.7	63.4	64.2	64.2	64.6	63.7	1.5	3.0	2.3	1.7	1.0	1.0	1.0	1.0	1.0	2.0	1.0	1.0
		0.6	64.3	64.4	64.6	64.0	64.6	64.5	65.1	64.4	3.3	3.3	3.0	3.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	66.5	65.7	66.0	65.9	66.5	65.7	66.0	65.9	3.7	4.0	4.7	2.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	0.2	62.3	62.3	62.4	62.5	63.0	63.0	63.8	63.3	2.0	3.3	4.3	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.7
		0.6	64.8	64.4	64.9	65.0	65.4	64.8	65.4	65.9	3.3	3.0	3.7	5.0	1.0	1.0	1.0	1.0	1.0	1.7	1.0	1.0
		1.0	66.3	67.0	67.3	68.0	66.3	67.0	67.3	68.0	3.0	3.0	3.3	3.7	1.0	1.0	1.0	1.0	1.0	1.0	1.7	1.0

APPENDIX A17

LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 2 (SET 1), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES.

Stage	Yield Gen. Architecture	H^2	Genotypic yield				Phenotypic yield			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	64.4 (±0.15)	64.7 (±0.15)	64.4 (±0.15)	64.2 (±0.15)	66.5 (±0.13)	66.4 (±0.14)	66.5 (±0.14)	66.0 (±0.13)
		0.6	63.9	64.5	65.0	65.0	65.0	65.6	66.2	66.0
		1.0	63.9	66.0	66.3	67.0	63.9	66.0	66.3	67.0
	uniform	0.2	62.2	62.1	62.0	62.0	63.5	63.4	63.1	63.2
		0.6	63.0	62.9	63.1	63.1	63.3	63.3	63.6	63.4
		1.0	65.0	64.7	64.8	64.7	65.0	64.7	64.8	64.7
	Negative Exponential	0.2	61.6	61.7	61.6	61.6	62.5	62.7	62.7	62.4
		0.6	63.1	63.2	63.3	62.8	63.8	63.8	63.9	63.4
		1.0	65.0	65.4	65.7	65.9	65.0	65.4	65.7	65.9
AYT-2	3-categories QTL	0.2	65.4 (±0.29)	65.8 (±0.29)	66.1 (±0.34)	65.3 (±0.36)	66.0 (±0.24)	66.9 (±0.26)	66.6 (±0.32)	66.2 (±0.27)
		0.6	65.5	66.1	67.0	67.0	66.0	66.7	67.7	67.5
		1.0	67.0	69.1	68.8	69.8	67.0	69.1	68.8	69.8
	uniform	0.2	62.9	62.8	62.8	62.9	63.7	63.4	63.4	63.5
		0.6	63.7	63.7	64.0	64.1	64.0	63.9	64.2	64.4
		1.0	66.3	66.2	66.0	65.9	66.3	66.2	66.0	65.9
	Negative Exponential	0.2	62.0	62.3	62.4	62.2	62.8	63.0	63.1	63.1
		0.6	64.5	64.2	64.2	63.8	64.9	64.5	64.6	64.1
		1.0	66.3	66.6	67.0	67.0	66.3	66.6	67.0	67.0
EYT	3-categories QTL	0.2	66.0 (±0.88)	66.2 (±0.35)	66.9 (±0.60)	65.3 (±0.80)	67.1 (±0.38)	67.4 (±0.33)	67.9 (±0.66)	67.0 (±0.71)
		0.6	66.2	66.9	67.8	68.0	67.4	67.9	68.3	68.9
		1.0	69.3	71.0	70.8	71.0	69.3	71.0	70.8	71.0
	uniform	0.2	63.4	63.4	62.8	63.4	64.0	63.7	64.2	64.0
		0.6	64.1	64.0	64.6	64.3	64.6	64.4	64.5	64.7
		1.0	66.6	66.7	66.8	67.1	66.6	66.7	66.8	67.1
	Negative Exponential	0.2	62.4	62.3	62.4	62.5	63.1	63.2	63.7	63.1
		0.6	65.0	64.4	64.4	65.0	65.5	65.1	65.0	65.0
		1.0	67.2	66.9	67.7	67.5	67.2	66.9	67.7	67.5

APPENDIX A18

LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 2 (SET 2), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES.

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	64.5 (±0.16)	64.8 (±0.15)	64.3 (±0.15)	63.8 (±0.15)	66.7 (±0.14)	66.5 (±0.13)	66.6 (±0.15)	65.7 (±0.13)
		0.6	63.9	64.5	65.4	65.4	65.0	65.5	66.5	66.5
		1.0	63.9	64.7	66.1	67.2	63.9	64.7	66.1	67.2
	uniform	0.2	62.0	62.1	62.1	62.1	63.4	63.4	63.3	63.2
		0.6	63.0	63.2	63.2	63.0	63.5	63.5	63.6	63.4
		1.0	65.2	64.7	65.0	64.8	65.2	64.7	65.0	64.8
	Negative Exponential	0.2	61.6	61.8	61.6	61.6	62.5	62.7	62.8	62.5
		0.6	63.3	63.0	63.1	62.6	63.8	63.7	63.6	63.1
		1.0	65.2	65.4	65.8	65.8	65.2	65.4	65.8	65.8
AYT-2	3-categories QTL	0.2	65.4 (±0.35)	66.0 (±0.29)	65.6 (±0.41)	65.5 (±0.25)	66.9 (±0.31)	66.7 (±0.25)	66.5 (±0.36)	66.4 (±0.21)
		0.6	65.4	66.7	67.0	67.5	66.1	67.3	67.4	67.9
		1.0	67.3	67.4	69.1	70.1	67.3	67.4	69.1	70.1
	uniform	0.2	62.9	62.8	62.9	62.8	63.6	63.6	63.6	63.4
		0.6	63.9	64.1	64.1	63.9	64.0	64.4	64.4	64.1
		1.0	66.5	65.7	66.1	65.9	66.5	65.7	66.1	65.9
	Negative Exponential	0.2	62.0	62.3	62.1	62.1	62.8	63.0	62.9	62.9
		0.6	64.5	63.9	64.4	63.4	65.0	64.3	64.8	63.7
		1.0	66.5	66.5	66.9	66.8	66.5	66.5	66.9	66.8
EYT	3-categories QTL	0.2	66.0 (±0.57)	66.4 (±0.43)	66.8 (±0.65)	66.0 (±0.22)	67.6 (±0.67)	66.6 (±0.53)	68.4 (±0.56)	66.4 (±0.41)
		0.6	66.2	67.9	68.7	68.0	67.4	68.5	69.7	68.1
		1.0	70.1	69.4	71.5	71.2	70.1	69.4	71.5	71.2
	uniform	0.2	63.4	62.9	63.0	63.7	64.4	63.8	64.0	64.0
		0.6	64.6	64.5	64.5	64.4	65.0	64.5	65.0	64.6
		1.0	66.8	66.3	66.5	66.3	66.8	66.3	66.5	66.3
	Negative Exponential	0.2	62.4	62.6	62.3	62.1	63.0	63.1	63.4	62.9
		0.6	65.2	64.3	64.4	64.2	65.7	64.8	65.0	64.3
		1.0	67.1	67.0	67.5	67.5	67.1	67.0	67.5	67.5

APPENDIX A19

LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 2 (SET 3), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES.

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	64.6 (±0.14)	64.7 (±0.14)	64.2 (±0.15)	63.9 (±0.15)	66.7 (±0.13)	66.7 (±0.13)	66.2 (±0.14)	66.0 (±0.13)
		0.6	64.2	64.6	65.3	65.3	65.2	65.5	66.5	66.4
		1.0	63.2	64.6	65.9	67.0	63.2	64.6	65.9	67.0
	uniform	0.2	62.1	62.1	62.2	62.3	63.3	63.4	63.4	63.6
		0.6	63.0	63.0	63.1	63.0	63.3	63.4	63.5	63.3
		1.0	65.2	64.5	64.5	64.6	65.2	64.5	64.5	64.6
	Negative Exponential	0.2	61.6	61.7	61.7	61.7	62.5	62.7	62.7	62.6
		0.6	63.2	63.1	63.0	62.8	63.7	63.6	63.7	63.4
		1.0	64.8	65.3	65.7	66.1	64.8	65.3	65.7	66.1
AYT-2	3-categories QTL	0.2	65.7 (±0.26)	65.8 (±0.22)	65.4 (±0.26)	65.3 (±0.31)	66.6 (±0.23)	66.9 (±0.19)	66.2 (±0.21)	65.9 (±0.26)
		0.6	65.8	66.0	67.4	67.5	66.6	66.7	68.2	67.9
		1.0	66.1	67.1	68.1	70.2	66.1	67.1	68.1	70.2
	uniform	0.2	62.9	62.9	62.9	63.1	63.5	63.8	63.8	63.9
		0.6	63.8	64.0	63.9	63.8	64.1	64.3	64.3	64.1
		1.0	66.2	65.3	65.6	65.5	66.2	65.3	65.6	65.5
	Negative Exponential	0.2	62.0	62.2	62.4	62.3	62.9	62.9	63.1	63.1
		0.6	64.4	64.1	64.1	63.8	64.9	64.5	64.5	64.2
		1.0	65.9	66.6	66.8	67.5	65.9	66.6	66.8	67.5
EYT	3-categories QTL	0.2	65.9 (±0.46)	66.0 (±0.32)	65.9 (±0.52)	65.9 (±0.67)	66.8 (±0.71)	66.8 (±0.20)	67.3 (±0.58)	66.1 (±0.55)
		0.6	66.6	67.1	67.6	68.5	67.6	67.5	68.2	69.5
		1.0	68.2	67.8	69.0	71.7	68.2	67.8	69.0	71.7
	uniform	0.2	63.4	63.1	63.7	63.4	64.2	64.2	64.6	63.7
		0.6	64.3	64.4	64.6	64.0	64.6	64.5	65.1	64.4
		1.0	66.5	65.7	66.0	65.9	66.5	65.7	66.0	65.9
	Negative Exponential	0.2	62.3	62.3	62.4	62.5	63.0	63.0	63.8	63.3
		0.6	64.8	64.4	64.9	65.0	65.4	64.8	65.4	65.9
		1.0	66.3	67.0	67.3	68.0	66.3	67.0	67.3	68.0

APPENDIX A20

DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY IN FOUNDER SET 1 USED IN EXPERIMENT 2 FOR NINE COMBINATIONS OF GENETIC ARCHITECTURE AND HERITABILITY FOR EACH OF FOUR STAGES IN SOYBEAN VARIETY DEVELOPMENT.

Stage	Yield genetic architecture	H^2	Accuracy	Sensitivity	Specificity
		0.20	0.88	0.41	0.93
	3-categories QTL	0.60	0.92	0.64	0.96
		1.00	1.00	1.00	1.00
		0.20	0.87	0.39	0.93
PYT	Uniform	0.60	0.93	0.68	0.96
		1.00	1.00	1.00	1.00
		0.20	0.87	0.39	0.93
	Negative Exponential	0.60	0.92	0.63	0.96
		1.00	1.00	1.00	1.00
		0.20	0.80	0.48	0.87
	3-categories QTL	0.60	0.84	0.58	0.90
		1.00	1.00	1.00	1.00
		0.20	0.82	0.53	0.89
AYT-1	Uniform	0.60	0.87	0.66	0.92
		1.00	1.00	1.00	1.00
		0.20	0.83	0.56	0.89
	Negative Exponential	0.60	0.85	0.62	0.91
		1.00	1.00	1.00	1.00
		0.20	0.88	0.55	0.93
	3-categories QTL	0.60	0.90	0.63	0.94
		1.00	1.00	1.00	1.00
		0.20	0.87	0.50	0.92
AYT-2	Uniform	0.60	0.91	0.65	0.95
		1.00	1.00	1.00	1.00
		0.20	0.86	0.48	0.92
	Negative Exponential	0.60	0.93	0.73	0.96
		1.00	1.00	1.00	1.00
		0.20	0.75	0.50	0.83
	3-categories QTL	0.60	0.75	0.50	0.83
		1.00	1.00	1.00	1.00
		0.20	0.80	0.60	0.87
EYT	Uniform	0.60	0.75	0.50	0.83
		1.00	1.00	1.00	1.00
		0.20	0.70	0.40	0.80
	Negative Exponential	0.60	0.75	0.50	0.83
		1.00	1.00	1.00	1.00

APPENDIX A21

DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY IN FOUNDER SET 2 USED IN EXPERIMENT 2 FOR NINE COMBINATIONS OF GENETIC ARCHITECTURE AND HERITABILITY FOR EACH OF FOUR STAGES IN SOYBEAN VARIETY DEVELOPMENT.

Stage	Yield genetic architecture	H^2	Accuracy	Sensitivity	Specificity
		0.20	0.87	0.38	0.93
	3-categories QTL	0.60	0.92	0.62	0.96
		1.00	1.00	1.00	1.00
		0.20	0.87	0.39	0.93
PYT	Uniform	0.60	0.93	0.67	0.96
		1.00	1.00	1.00	1.00
		0.20	0.87	0.40	0.93
	Negative Exponential	0.60	0.92	0.62	0.96
		1.00	1.00	1.00	1.00
		0.20	0.80	0.49	0.88
	3-categories QTL	0.60	0.85	0.60	0.90
		1.00	1.00	1.00	1.00
		0.20	0.80	0.50	0.88
AYT-1	Uniform	0.60	0.86	0.65	0.92
		1.00	1.00	1.00	1.00
		0.20	0.83	0.56	0.89
	Negative Exponential	0.60	0.88	0.69	0.92
		1.00	1.00	1.00	1.00
		0.20	0.87	0.53	0.93
	3-categories QTL	0.60	0.89	0.60	0.94
		1.00	1.00	1.00	1.00
		0.20	0.87	0.50	0.92
AYT-2	Uniform	0.60	0.92	0.70	0.95
		1.00	1.00	1.00	1.00
		0.20	0.83	0.38	0.90
	Negative Exponential	0.60	0.92	0.70	0.95
		1.00	1.00	1.00	1.00
		0.20	0.70	0.40	0.80
	3-categories QTL	0.60	0.80	0.60	0.87
		1.00	1.00	1.00	1.00
		0.20	0.80	0.60	0.87
EYT	Uniform	0.60	0.80	0.60	0.87
		1.00	1.00	1.00	1.00
		0.20	0.70	0.40	0.80
	Negative Exponential	0.60	0.80	0.60	0.87
		1.00	1.00	1.00	1.00

APPENDIX A22

DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY IN FOUNDER SET 3 USED IN EXPERIMENT 2 FOR NINE COMBINATIONS OF GENETIC ARCHITECTURE AND HERITABILITY FOR EACH OF FOUR STAGES IN SOYBEAN VARIETY DEVELOPMENT.

Stage	Yield genetic architecture	H^2	Accuracy	Sensitivity	Specificity
PYT	3-categories QTL	0.20	0.87	0.39	0.93
		0.60	0.92	0.64	0.96
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.87	0.40	0.93
		0.60	0.93	0.67	0.96
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.87	0.40	0.93
		0.60	0.92	0.63	0.96
		1.00	1.00	1.00	1.00
AYT-1	3-categories QTL	0.20	0.79	0.46	0.87
		0.60	0.85	0.61	0.90
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.81	0.52	0.88
		0.60	0.89	0.72	0.93
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.83	0.56	0.89
		0.60	0.87	0.66	0.92
		1.00	1.00	1.00	1.00
AYT-2	3-categories QTL	0.20	0.86	0.48	0.92
		0.60	0.91	0.65	0.95
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.87	0.50	0.92
		0.60	0.91	0.65	0.95
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.85	0.45	0.92
		0.60	0.93	0.73	0.96
		1.00	1.00	1.00	1.00
EYT	3-categories QTL	0.20	0.75	0.50	0.83
		0.60	0.80	0.60	0.87
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.80	0.60	0.87
		0.60	0.85	0.70	0.90
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.75	0.50	0.83
		0.60	0.85	0.70	0.90
		1.00	1.00	1.00	1.00

APPENDIX A23

ANALYSIS OF VARIANCE FOR SIMULATED GENOTYPIC VALUES FOR YIELD DUE TO FOUNDER SETS, GENETIC ARCHITECTURES AND HERITABILITY OF EYT LINES IN EXPERIMENT 3, IN WHICH FOUNDER CROSSES WERE SELECTED BASED ON TRAINING SETS CONSISTING OF 1,000 SNP ASSAYS AND SIMULATED YIELDS IN 112 FOUNDERS PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD.

Source	**Sum of Squares**	**Mean Square**	**F Ratio**	**Prob > F**
Model	25415.3	317.6	346.7	<.0001*
Error	8584.5	0.9		
C. Total	33999.9			
Source		**S.Squares**	**F Ratio**	**Prob > F**
Founder_Set		6.1	3.3	0.0350*
Stage		2245.4	1225.3	<.0001*
Founder_Set*Stage		0.7	0.1	0.9401
Genetic_architecture		2574.7	1404.9	<.0001*
Stage* Genetic_architecture		324.3	88.4	<.0001*
Founder_Set* Genetic_architecture		13.6	3.7	0.0049*
Founder_Set*Stage* Genetic_architecture		0.9	0.1	0.9982
Heritability		4193.1	2288.1	<.0001*
Stage* Heritability		193.5	52.7	<.0001*
Founder_Set* Heritability		5.8	1.5	0.1750
Founder_Set*Stage* Heritability		1.3	0.1	0.9938
Genetic_architecture* Heritability		130.8	35.7	<.0001*
Stage* Genetic_architecture* Heritability		91.1	12.4	<.0001*
Founder_Set* Genetic_architecture* Heritability		16.5	2.2	0.0205*
Founder_Set*Stage* Genetic_archit.* Heritability		3.1	0.2	0.9996

APPENDIX A24

SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FROM EXPERIMENT 3 FOR YA SET.

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield				Lodging				SCN				PRR			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	64.2	64.4	64.5	64.1	66.4	66.6	66.7	66.3	2.9	2.9	2.7	3.1	1.0	1.0	1.0	1.0	1.1	1.2	1.2	1.2
		0.6	64.3	65.1	65.8	66.4	65.5	66.2	67.0	67.5	3.3	3.0	3.1	3.3	1.0	1.0	1.0	1.0	1.1	1.3	1.3	1.2
		1.0	64.4	65.6	66.6	66.8	64.4	65.6	66.6	66.8	3.0	2.9	2.9	2.6	1.0	1.0	1.0	1.0	1.1	1.2	1.1	1.1
	uniform	0.2	61.9	62.0	62.2	62.1	63.2	63.3	63.3	63.3	3.1	2.7	3.0	3.1	1.0	1.0	1.0	1.0	1.3	1.1	1.3	1.2
		0.6	62.7	62.9	63.0	63.0	63.2	63.3	63.4	63.4	3.3	3.3	3.2	3.6	1.0	1.0	1.0	1.0	1.1	1.2	1.3	1.1
		1.0	65.4	64.8	65.1	64.6	65.4	64.8	65.1	64.6	3.3	2.9	3.2	3.4	1.0	1.0	1.0	1.0	1.1	1.0	1.1	1.1
	Negative Exponential	0.2	61.6	61.8	61.7	61.6	62.6	62.6	62.7	62.5	3.1	2.9	3.2	3.2	1.0	1.0	1.0	1.0	1.2	1.2	1.2	1.2
		0.6	63.4	63.5	63.4	62.8	64.0	64.0	64.0	63.5	2.8	3.0	3.3	3.1	1.0	1.0	1.0	1.0	1.2	1.2	1.2	1.2
		1.0	65.2	65.6	65.7	65.4	65.2	65.6	65.7	65.4	3.2	3.2	3.0	3.0	1.0	1.0	1.0	1.0	1.1	1.2	1.3	1.1
AYT-2	3-categories QTL	0.2	65.7	65.3	65.8	64.9	67.1	66.6	67.1	66.1	2.5	3.4	3.0	2.9	1.0	1.0	1.0	1.0	1.4	1.0	1.2	1.2
		0.6	66.4	66.8	67.0	68.5	66.9	67.7	67.7	69.4	3.5	3.4	3.4	2.8	1.0	1.0	1.0	1.0	1.6	1.0	1.6	1.6
		1.0	67.3	68.3	69.6	69.7	67.3	68.3	69.6	69.7	3.4	2.8	2.4	3.2	1.0	1.0	1.0	1.0	1.0	1.0	1.2	1.2
	uniform	0.2	62.2	62.8	63.2	62.8	63.0	63.7	63.7	63.3	4.2	2.6	2.9	2.6	1.0	1.0	1.0	1.0	1.0	1.2	1.4	1.0
		0.6	63.6	63.8	63.9	63.8	63.7	64.1	64.1	64.0	3.3	3.5	3.6	3.8	1.0	1.0	1.0	1.0	1.0	1.2	1.4	1.0
		1.0	66.7	65.6	66.1	65.5	66.7	65.6	66.1	65.5	3.6	3.3	2.8	3.7	1.0	1.0	1.0	1.0	1.0	1.0	1.2	1.2
	Negative Exponential	0.2	62.1	62.3	62.4	62.1	62.8	63.1	63.3	62.8	2.8	2.6	3.1	3.0	1.0	1.0	1.0	1.0	1.4	1.4	1.2	1.4
		0.6	64.4	64.4	64.4	63.6	64.9	64.7	64.8	64.0	3.1	3.1	3.3	2.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	66.6	66.5	66.8	66.7	66.6	66.5	66.8	66.7	3.0	3.8	3.5	3.2	1.0	1.0	1.0	1.0	1.0	1.0	1.2	1.0
EYT	3-categories QTL	0.2	65.7	65.6	66.3	64.9	67.7	67.2	67.3	67.6	2.7	3.0	2.7	1.5	1.0	1.0	1.0	1.0	1.7	1.0	1.0	2.0
		0.6	67.7	67.6	67.4	68.4	68.0	67.9	68.3	70.1	4.5	4.5	3.3	3.3	1.0	1.0	1.0	1.0	1.0	1.0	2.3	2.3
		1.0	68.2	69.3	71.2	71.3	68.2	69.3	71.2	71.3	4.0	2.7	1.7	3.3	1.0	1.0	1.0	1.0	1.0	1.0	1.7	1.0
	uniform	0.2	62.0	63.2	63.4	63.0	63.4	63.6	63.8	63.9	3.0	3.0	4.0	1.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	63.9	64.0	64.3	64.0	64.2	64.4	64.8	64.3	3.5	3.0	3.7	3.5	1.0	1.0	1.0	1.0	1.0	1.0	1.7	1.0
		1.0	67.1	66.1	66.5	66.0	67.1	66.1	66.5	66.0	4.3	5.0	2.7	4.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	2.0
	Negative Exponential	0.2	62.6	62.6	62.4	62.0	63.1	62.9	62.8	62.6	2.0	3.5	3.0	3.3	1.0	1.0	1.0	1.0	2.0	1.0	1.0	1.7
		0.6	64.8	64.6	65.3	65.0	65.3	65.0	65.6	65.0	3.0	3.3	3.7	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	67.2	66.9	67.7	67.7	67.2	66.9	67.7	67.7	2.3	3.5	4.0	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.7	1.0

APPENDIX A25

SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FROM EXPERIMENT 3 FOR YB SET.

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield				Lodging				SCN				PRR			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	64.0	64.7	64.6	64.2	66.1	66.8	66.5	66.2	2.7	3.0	2.8	3.2	1.0	1.0	1.0	1.0	1.1	1.3	1.2	1.1
		0.6	64.1	64.7	65.9	65.8	65.1	65.7	67.2	67.1	3.3	3.4	3.3	3.3	1.0	1.0	1.0	1.0	1.2	1.2	1.2	1.2
		1.0	64.4	65.6	66.8	67.5	64.4	65.6	66.8	67.5	2.9	2.9	3.1	3.3	1.0	1.0	1.0	1.0	1.2	1.0	1.1	1.1
	uniform	0.2	62.3	62.3	62.3	62.3	63.6	63.6	63.5	63.3	2.9	2.9	3.0	3.2	1.0	1.0	1.0	1.0	1.3	1.1	1.2	1.2
		0.6	63.0	63.0	63.4	63.6	63.5	63.5	63.9	64.0	3.1	3.0	3.1	3.7	1.0	1.0	1.0	1.0	1.1	1.1	1.1	1.1
		1.0	65.5	65.0	65.4	65.1	65.5	65.0	65.4	65.1	3.1	2.9	3.1	3.5	1.0	1.0	1.0	1.0	1.2	1.2	1.2	1.2
	Negative Exponential	0.2	61.6	61.8	61.7	61.6	62.6	62.7	62.7	62.5	3.2	3.3	3.5	3.1	1.0	1.0	1.0	1.0	1.1	1.1	1.1	1.1
		0.6	63.2	63.1	63.4	63.0	63.7	63.7	63.9	63.6	3.1	3.2	3.3	3.1	1.0	1.0	1.0	1.0	1.1	1.2	1.3	1.3
		1.0	65.1	65.4	65.8	65.8	65.1	65.4	65.8	65.8	3.3	3.3	3.1	3.5	1.0	1.0	1.0	1.0	1.1	1.3	1.3	1.3
AYT-2	3-categories QTL	0.2	65.8	65.7	65.7	65.3	66.3	66.9	66.8	66.6	2.8	2.5	2.9	3.3	1.0	1.0	1.0	1.0	1.2	1.6	1.0	1.2
		0.6	65.8	66.2	67.8	67.6	66.7	67.2	68.4	68.1	3.2	3.8	3.5	3.6	1.0	1.0	1.0	1.0	1.2	1.4	1.2	1.2
		1.0	67.6	68.5	69.3	70.0	67.6	68.5	69.3	70.0	2.2	2.9	2.8	2.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.4
	uniform	0.2	63.0	63.2	63.1	62.8	63.8	63.9	63.6	63.4	2.6	3.1	2.8	3.4	1.0	1.0	1.0	1.0	1.2	1.2	1.6	1.2
		0.6	63.7	63.6	64.0	64.4	63.8	64.0	64.5	64.7	2.7	3.4	3.2	3.6	1.0	1.0	1.0	1.0	1.0	1.0	1.2	1.0
		1.0	66.5	66.5	66.3	66.2	66.5	66.5	66.3	66.2	3.5	3.1	3.5	3.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	0.2	62.2	62.4	62.7	62.0	63.0	62.9	63.5	62.9	3.1	3.2	3.0	3.3	1.0	1.0	1.0	1.0	1.0	1.0	1.2	1.0
		0.6	64.1	64.3	64.1	64.0	64.2	64.7	64.5	64.2	3.1	2.9	3.0	2.9	1.0	1.0	1.0	1.0	1.2	1.6	1.4	1.0
		1.0	66.1	66.5	66.8	67.3	66.1	66.5	66.8	67.3	2.8	3.5	3.5	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.4	1.4
EYT	3-categories QTL	0.2	66.2	66.2	66.3	65.5	67.4	67.1	67.0	66.3	2.7	4.0	2.5	4.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	66.6	66.3	69.4	68.2	68.7	67.2	69.4	69.6	4.0	4.7	3.0	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.7	1.0
		1.0	69.7	69.6	70.6	70.8	69.7	69.6	70.6	70.8	2.0	2.0	2.0	1.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.7
	uniform	0.2	63.2	63.6	64.2	62.9	63.9	64.5	64.1	63.8	3.0	1.7	2.5	3.3	1.0	1.0	1.0	1.0	1.0	1.0	2.0	1.7
		0.6	63.7	64.1	64.4	64.9	64.4	64.4	65.0	65.4	4.0	3.7	3.0	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	66.8	67.2	67.6	66.8	66.8	67.2	67.6	66.8	3.0	2.7	2.5	3.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	0.2	62.3	62.6	62.8	62.3	63.4	62.9	63.1	63.5	2.7	3.7	2.0	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	64.7	64.9	64.6	65.1	65.1	65.2	65.1	65.3	2.3	4.5	3.3	3.5	1.0	1.0	1.0	1.0	1.0	2.0	1.7	1.0
		1.0	66.3	67.2	67.2	68.1	66.3	67.2	67.2	68.1	2.0	3.0	3.7	4.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.7

APPENDIX A26

SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FROM EXPERIMENT 3 FOR YC SET.

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield				Lodging				SCN				PRR			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	64.2	64.1	64.2	64.0	66.4	66.5	66.6	65.8	2.9	3.1	3.0	3.2	1.0	1.0	1.0	1.0	1.2	1.3	1.2	1.1
		0.6	64.0	64.3	65.3	65.4	65.0	65.4	66.4	66.7	3.5	3.3	3.1	3.3	1.0	1.0	1.0	1.0	1.1	1.2	1.2	1.3
		1.0	64.7	66.1	66.9	67.4	64.7	66.1	66.9	67.4	2.9	3.0	3.1	2.9	1.0	1.0	1.0	1.0	1.1	1.2	1.1	1.2
	uniform	0.2	62.2	62.3	62.3	62.6	63.6	63.4	63.6	63.6	3.2	2.9	2.8	2.8	1.0	1.0	1.0	1.0	1.3	1.2	1.3	1.2
		0.6	62.9	62.9	63.3	63.2	63.3	63.4	63.7	63.7	3.2	3.1	3.2	3.7	1.0	1.0	1.0	1.0	1.2	1.1	1.3	1.1
		1.0	65.0	65.0	65.3	64.9	65.0	65.0	65.3	64.9	3.5	3.5	3.2	3.4	1.0	1.0	1.0	1.0	1.2	1.3	1.1	1.2
	Negative Exponential	0.2	61.5	61.6	61.6	61.7	62.3	62.5	62.7	62.6	3.0	3.1	3.4	3.2	1.0	1.0	1.0	1.0	1.2	1.2	1.2	1.1
		0.6	63.6	63.4	63.4	63.3	64.2	64.0	64.1	63.8	3.1	3.0	3.2	2.9	1.0	1.0	1.0	1.0	1.3	1.3	1.2	1.2
		1.0	65.2	65.6	65.8	66.3	65.2	65.6	65.8	66.3	3.3	3.0	3.1	3.3	1.0	1.0	1.0	1.0	1.1	1.3	1.1	1.3
AYT-2	3-categories QTL	0.2	65.3	65.2	65.9	65.1	66.8	66.6	67.2	66.5	2.7	3.2	2.5	3.3	1.0	1.0	1.0	1.0	1.0	1.4	1.4	1.2
		0.6	66.0	66.0	66.5	67.5	66.9	66.9	67.3	68.0	3.6	3.7	2.9	3.4	1.0	1.0	1.0	1.0	1.2	1.2	1.0	1.2
		1.0	67.7	69.4	69.6	69.6	67.7	69.4	69.6	69.6	2.6	2.6	2.7	3.4	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.6
	uniform	0.2	63.0	63.3	63.2	63.3	63.7	64.1	63.8	64.1	3.5	2.5	2.4	2.7	1.0	1.0	1.0	1.0	1.6	1.4	1.2	1.2
		0.6	64.1	63.9	63.9	64.1	64.2	64.1	64.4	64.4	2.8	3.4	2.6	4.2	1.0	1.0	1.0	1.0	1.0	1.0	1.2	1.2
		1.0	66.0	66.3	66.2	66.2	66.0	66.3	66.2	66.2	3.9	3.4	3.4	3.5	1.0	1.0	1.0	1.0	1.2	1.2	1.0	1.0
	Negative Exponential	0.2	62.0	62.1	62.3	62.4	62.7	62.8	63.2	63.1	2.4	2.8	3.8	3.3	1.0	1.0	1.0	1.0	1.0	1.0	1.2	1.2
		0.6	64.4	64.4	64.3	64.5	64.8	64.8	64.9	64.8	3.5	2.9	3.2	2.9	1.0	1.0	1.0	1.0	1.2	1.8	1.4	1.2
		1.0	66.2	66.8	67.0	67.6	66.2	66.8	67.0	67.6	3.1	2.9	2.6	3.4	1.0	1.0	1.0	1.0	1.2	1.2	1.0	1.2
EYT	3-categories QTL	0.2	66.0	65.8	66.4	65.3	67.2	67.9	66.1	67.7	3.0	3.7	2.0	4.0	1.0	1.0	1.0	1.0	1.0	1.0	1.7	1.0
		0.6	67.6	66.8	66.6	68.0	68.4	67.7	68.0	68.1	4.0	4.0	3.7	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.5
		1.0	68.3	71.0	70.9	70.5	68.3	71.0	70.9	70.5	2.0	2.3	3.0	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	0.2	63.3	64.0	63.4	63.6	64.1	64.9	64.5	64.3	4.0	2.3	3.0	2.7	1.0	1.0	1.0	1.0	1.0	1.7	1.0	1.0
		0.6	65.1	64.3	63.8	64.8	65.1	64.6	64.2	64.7	2.0	3.3	1.0	3.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	66.6	67.0	66.7	66.6	66.6	67.0	66.7	66.6	2.0	3.3	4.0	3.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	0.2	62.4	62.3	62.6	62.7	63.1	62.9	63.4	63.1	2.0	3.5	4.3	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	65.3	64.5	65.0	65.1	65.2	65.1	65.5	65.2	3.0	3.0	3.3	3.3	1.0	1.0	1.0	1.0	1.0	1.7	1.0	1.0
		1.0	67.4	67.1	67.7	68.0	67.4	67.1	67.7	68.0	4.0	2.3	2.3	3.7	1.0	1.0	1.0	1.0	1.0	1.7	1.0	1.0

APPENDIX A27

LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 3 (YA SET), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES.

Stage	Yield Gen. Architecture	H^2	Genotypic yield				Phenotypic yield			
			1.8-2.4	**2.4-3.0**	**3.0-3.6**	**3.6-4.2**	**1.8-2.4**	**2.4-3.0**	**3.0-3.6**	**3.6-4.2**
AYT-1	**3-categories QTL**	**0.2**	64.2 (±0.10)	64.4 (±0.10)	64.5 (±0.10)	64.1 (±0.10)	66.4 (±0.09)	66.6 (±0.09)	66.7 (±0.10)	66.3 (±0.09)
		0.6	64.3	65.1	65.8	66.4	65.5	66.2	67.0	67.5
		1.0	64.4	65.6	66.6	66.8	64.4	65.6	66.6	66.8
	uniform	**0.2**	61.9	62.0	62.2	62.1	63.2	63.3	63.3	63.3
		0.6	62.7	62.9	63.0	63.0	63.2	63.3	63.4	63.4
		1.0	65.4	64.8	65.1	64.6	65.4	64.8	65.1	64.6
	Negative Exponential	**0.2**	61.6	61.8	61.7	61.6	62.6	62.6	62.7	62.5
		0.6	63.4	63.5	63.4	62.8	64.0	64.0	64.0	63.5
		1.0	65.2	65.6	65.7	65.4	65.2	65.6	65.7	65.4
AYT-2	**3-categories QTL**	**0.2**	65.7 (±0.16)	65.3 (±0.15)	65.8 (±0.23)	64.9 (±0.22)	67.1 (±0.15)	66.6 (±0.14)	67.1 (±0.21)	66.1 (±0.20)
		0.6	66.4	66.8	67.0	68.5	66.9	67.7	67.7	69.4
		1.0	67.3	68.3	69.6	69.7	67.3	68.3	69.6	69.7
	uniform	**0.2**	62.2	62.8	63.2	62.8	63.0	63.7	63.7	63.3
		0.6	63.6	63.8	63.9	63.8	63.7	64.1	64.1	64.0
		1.0	66.7	65.6	66.1	65.5	66.7	65.6	66.1	65.5
	Negative Exponential	**0.2**	62.1	62.3	62.4	62.1	62.8	63.1	63.3	62.8
		0.6	64.4	64.4	64.4	63.6	64.9	64.7	64.8	64.0
		1.0	66.6	66.5	66.8	66.7	66.6	66.5	66.8	66.7
EYT	**3-categories QTL**	**0.2**	65.7 (±0.33)	65.6 (±0.36)	66.3 (±0.41)	64.9 (±0.54)	67.7 (±0.28)	67.2 (±0.32)	67.3 (±0.41)	67.6 (±0.47)
		0.6	67.7	67.6	67.4	68.4	68.0	67.9	68.3	70.1
		1.0	68.2	69.3	71.2	71.3	68.2	69.3	71.2	71.3
	uniform	**0.2**	62.0	63.2	63.4	63.0	63.4	63.6	63.8	63.9
		0.6	63.9	64.0	64.3	64.0	64.2	64.4	64.8	64.3
		1.0	67.1	66.1	66.5	66.0	67.1	66.1	66.5	66.0
	Negative Exponential	**0.2**	62.6	62.6	62.4	62.0	63.1	62.9	62.8	62.6
		0.6	64.8	64.6	65.3	65.0	65.3	65.0	65.6	65.0
		1.0	67.2	66.9	67.7	67.7	67.2	66.9	67.7	67.7

APPENDIX A28

LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 3 (YB SET), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES.

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	64.0 (±0.11)	64.7 (±0.10)	64.6 (±0.10)	64.2 (±0.10)	66.1 (±0.10)	66.8 (±0.10)	66.5 (±0.10)	66.2 (±0.09)
		0.6	64.1	64.7	65.9	65.8	65.1	65.7	67.2	67.1
		1.0	64.4	65.6	66.8	67.5	64.4	65.6	66.8	67.5
	uniform	0.2	62.3	62.3	62.3	62.3	63.6	63.6	63.5	63.3
		0.6	63.0	63.0	63.4	63.6	63.5	63.5	63.9	64.0
		1.0	65.5	65.0	65.4	65.1	65.5	65.0	65.4	65.1
	Negative Exponential	0.2	61.6	61.8	61.7	61.6	62.6	62.7	62.7	62.5
		0.6	63.2	63.1	63.4	63.0	63.7	63.7	63.9	63.6
		1.0	65.1	65.4	65.8	65.8	65.1	65.4	65.8	65.8
AYT-2	3-categories QTL	0.2	65.8 (±0.18)	65.7 (±0.18)	65.7 (±0.23)	65.3 (±0.18)	66.3 (±0.15)	66.9 (±0.17)	66.8 (±0.20)	66.6 (±0.16)
		0.6	65.8	66.2	67.8	67.6	66.7	67.2	68.4	68.1
		1.0	67.6	68.5	69.3	70.0	67.6	68.5	69.3	70.0
	uniform	0.2	63.0	63.2	63.1	62.8	63.8	63.9	63.6	63.4
		0.6	63.7	63.6	64.0	64.4	63.8	64.0	64.5	64.7
		1.0	66.5	66.5	66.3	66.2	66.5	66.5	66.3	66.2
	Negative Exponential	0.2	62.2	62.4	62.7	62.0	63.0	62.9	63.5	62.9
		0.6	64.1	64.3	64.1	64.0	64.2	64.7	64.5	64.2
		1.0	66.1	66.5	66.8	67.3	66.1	66.5	66.8	67.3
EYT	3-categories QTL	0.2	66.2 (±0.36)	66.2 (±0.26)	66.3 (±0.46)	65.5 (±0.31)	67.4 (±0.25)	67.1 (±0.30)	67.0 (±0.48)	66.3 (±0.31)
		0.6	66.6	66.3	69.4	68.2	68.7	67.2	69.4	69.6
		1.0	69.7	69.6	70.6	70.8	69.7	69.6	70.6	70.8
	uniform	0.2	63.2	63.6	64.2	62.9	63.9	64.5	64.1	63.8
		0.6	63.7	64.1	64.4	64.9	64.4	64.4	65.0	65.4
		1.0	66.8	67.2	67.6	66.8	66.8	67.2	67.6	66.8
	Negative Exponential	0.2	62.3	62.6	62.8	62.3	63.4	62.9	63.1	63.5
		0.6	64.7	64.9	64.6	65.1	65.1	65.2	65.1	65.3
		1.0	66.3	67.2	67.2	68.1	66.3	67.2	67.2	68.1

APPENDIX A29

LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 3 (YC SET), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES.

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield			
			1.8-2.4	**2.4-3.0**	**3.0-3.6**	**3.6-4.2**	**1.8-2.4**	**2.4-3.0**	**3.0-3.6**	**3.6-4.2**
AYT-1	**3-categories QTL**	**0.2**	64.2 (±0.10)	64.1 (±0.10)	64.2 (±0.10)	64.0 (±0.10)	66.4 (±0.10)	66.5 (±0.09)	66.6 (±0.09)	65.8 (±0.09)
		0.6	64.0	64.3	65.3	65.4	65.0	65.4	66.4	66.7
		1.0	64.7	66.1	66.9	67.4	64.7	66.1	66.9	67.4
	uniform	**0.2**	62.2	62.3	62.3	62.6	63.6	63.4	63.6	63.6
		0.6	62.9	62.9	63.3	63.2	63.3	63.4	63.7	63.7
		1.0	65.0	65.0	65.3	64.9	65.0	65.0	65.3	64.9
	Negative Exponential	**0.2**	61.5	61.6	61.6	61.7	62.3	62.5	62.7	62.6
		0.6	63.6	63.4	63.4	63.3	64.2	64.0	64.1	63.8
		1.0	65.2	65.6	65.8	66.3	65.2	65.6	65.8	66.3
AYT-2	**3-categories QTL**	**0.2**	65.3 (±0.20)	65.2 (±0.22)	65.9 (±0.18)	65.1 (±0.18)	66.8 (±0.18)	66.6 (±0.20)	67.2 (±0.18)	66.5 (±0.19)
		0.6	66.0	66.0	66.5	67.5	66.9	66.9	67.3	68.0
		1.0	67.7	69.4	69.6	69.6	67.7	69.4	69.6	69.6
	uniform	**0.2**	63.0	63.3	63.2	63.3	63.7	64.1	63.8	64.1
		0.6	64.1	63.9	63.9	64.1	64.2	64.1	64.4	64.4
		1.0	66.0	66.3	66.2	66.2	66.0	66.3	66.2	66.2
	Negative Exponential	**0.2**	62.0	62.1	62.3	62.4	62.7	62.8	63.2	63.1
		0.6	64.4	64.4	64.3	64.5	64.8	64.8	64.9	64.8
		1.0	66.2	66.8	67.0	67.6	66.2	66.8	67.0	67.6
EYT	**3-categories QTL**	**0.2**	66.0 (±0.74)	65.8 (±0.28)	66.4 (±0.30)	65.3 (±0.34)	67.2 (±0.81)	67.9 (±0.22)	66.1 (±0.28)	67.7 (±0.32)
		0.6	67.6	66.8	66.6	68.0	68.4	67.7	68.0	68.1
		1.0	68.3	71.0	70.9	70.5	68.3	71.0	70.9	70.5
	uniform	**0.2**	63.3	64.0	63.4	63.6	64.1	64.9	64.5	64.3
		0.6	65.1	64.3	63.8	64.8	65.1	64.6	64.2	64.7
		1.0	66.6	67.0	66.7	66.6	66.6	67.0	66.7	66.6
	Negative Exponential	**0.2**	62.4	62.3	62.6	62.7	63.1	62.9	63.4	63.1
		0.6	65.3	64.5	65.0	65.1	65.2	65.1	65.5	65.2
		1.0	67.4	67.1	67.7	68.0	67.4	67.1	67.7	68.0

APPENDIX A30

DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY METRICS FOR YA SET OF FOUNDER CROSSES FOR EACH HERITABILITY GENETIC ARCHITECTURE COMBINATION AND STAGE OF DEVELOPMENT IN SOYBEAN VARIETY DEVELOPMENT. EXPERIMENT 3 CONSISTS OF FOUNDER CROSSES THAT WERE SELECTED BASED ON TRAINING SETS CONSISTING OF 1,000 SNP ASSAYS AND SIMULATED YIELDS IN 112 FOUNDERS PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD.

Stage	Yield genetic architecture	H^2	Accuracy	Sensitivity	Specificity
		0.20	0.87	0.38	0.93
	3-categories QTL	0.60	0.92	0.62	0.96
		1.00	1.00	1.00	1.00
		0.20	0.88	0.39	0.93
PYT	Uniform	0.60	0.94	0.68	0.96
		1.00	1.00	1.00	1.00
		0.20	0.88	0.39	0.93
	Negative Exponential	0.60	0.92	0.62	0.96
		1.00	1.00	1.00	1.00
		0.20	0.79	0.45	0.87
	3-categories QTL	0.60	0.85	0.62	0.91
		1.00	1.00	1.00	1.00
		0.20	0.81	0.51	0.88
AYT-1	Uniform	0.60	0.89	0.71	0.93
		1.00	1.00	1.00	1.00
		0.20	0.82	0.55	0.89
	Negative Exponential	0.60	0.87	0.66	0.92
		1.00	1.00	1.00	1.00
		0.20	0.84	0.40	0.91
	3-categories QTL	0.60	0.94	0.78	0.97
		1.00	1.00	1.00	1.00
		0.20	0.83	0.38	0.90
AYT-2	Uniform	0.60	0.91	0.65	0.95
		1.00	1.00	1.00	1.00
		0.20	0.86	0.48	0.92
	Negative Exponential	0.60	0.91	0.68	0.95
		1.00	1.00	1.00	1.00
		0.20	0.70	0.40	0.80
	3-categories QTL	0.60	0.70	0.40	0.80
		1.00	1.00	1.00	1.00
		0.20	0.70	0.40	0.80
EYT	Uniform	0.60	0.75	0.50	0.83
		1.00	1.00	1.00	1.00
		0.20	0.70	0.40	0.80
	Negative Exponential	0.60	0.90	0.80	0.93
		1.00	1.00	1.00	1.00

APPENDIX A31

DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY METRICS FOR YB SET OF FOUNDER CROSSES FOR EACH HERITABILITY GENETIC ARCHITECTURE COMBINATION AND STAGE OF DEVELOPMENT IN SOYBEAN VARIETY DEVELOPMENT. EXPERIMENT 3 CONSISTS OF FOUNDER CROSSES THAT WERE SELECTED BASED ON TRAINING SETS CONSISTING OF 1,000 SNP ASSAYS AND SIMULATED YIELDS IN 112 FOUNDERS PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD.

Stage	Yield genetic architecture	H^2	Accuracy	Sensitivity	Specificity
		0.20	0.88	0.39	0.93
	3-categories QTL	0.60	0.93	0.63	0.96
		1.00	1.00	1.00	1.00
		0.20	0.87	0.37	0.93
PYT	Uniform	0.60	0.93	0.63	0.96
		1.00	1.00	1.00	1.00
		0.20	0.88	0.38	0.93
	Negative Exponential	0.60	0.92	0.62	0.96
		1.00	1.00	1.00	1.00
		0.20	0.80	0.48	0.88
	3-categories QTL	0.60	0.85	0.61	0.91
		1.00	1.00	1.00	1.00
		0.20	0.82	0.54	0.89
AYT-1	Uniform	0.60	0.86	0.65	0.92
		1.00	1.00	1.00	1.00
		0.20	0.82	0.55	0.89
	Negative Exponential	0.60	0.85	0.62	0.91
		1.00	1.00	1.00	1.00
		0.20	0.87	0.50	0.92
	3-categories QTL	0.60	0.91	0.68	0.95
		1.00	1.00	1.00	1.00
		0.20	0.85	0.45	0.92
AYT-2	Uniform	0.60	0.88	0.55	0.93
		1.00	1.00	1.00	1.00
		0.20	0.86	0.48	0.92
	Negative Exponential	0.60	0.90	0.63	0.94
		1.00	1.00	1.00	1.00
		0.20	0.80	0.60	0.87
	3-categories QTL	0.60	0.80	0.60	0.87
		1.00	1.00	1.00	1.00
		0.20	0.80	0.60	0.87
EYT	Uniform	0.60	0.85	0.70	0.90
		1.00	1.00	1.00	1.00
		0.20	0.70	0.40	0.80
	Negative Exponential	0.60	0.90	0.80	0.93
		1.00	1.00	1.00	1.00

APPENDIX A32

DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY METRICS FOR YC SET OF FOUNDER CROSSES FOR EACH HERITABILITY GENETIC ARCHITECTURE COMBINATION AND STAGE OF DEVELOPMENT IN SOYBEAN VARIETY DEVELOPMENT. EXPERIMENT 3 CONSISTS OF FOUNDER CROSSES THAT WERE SELECTED BASED ON TRAINING SETS CONSISTING OF 1,000 SNP ASSAYS AND SIMULATED YIELDS IN 112 FOUNDERS PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD.

Stage	Yield genetic architecture	H^2	Accuracy	Sensitivity	Specificity
		0.20	0.87	0.37	0.93
	3-categories QTL	0.60	0.92	0.62	0.96
		1.00	1.00	1.00	1.00
		0.20	0.88	0.40	0.93
PYT	Uniform	0.60	0.94	0.68	0.96
		1.00	1.00	1.00	1.00
		0.20	0.87	0.37	0.93
	Negative Exponential	0.60	0.92	0.59	0.95
		1.00	1.00	1.00	1.00
		0.20	0.78	0.44	0.86
	3-categories QTL	0.60	0.85	0.61	0.91
		1.00	1.00	1.00	1.00
		0.20	0.83	0.57	0.90
AYT-1	Uniform	0.60	0.88	0.68	0.92
		1.00	1.00	1.00	1.00
		0.20	0.82	0.54	0.89
	Negative Exponential	0.60	0.86	0.65	0.91
		1.00	1.00	1.00	1.00
		0.20	0.85	0.43	0.91
	3-categories QTL	0.60	0.91	0.68	0.95
		1.00	1.00	1.00	1.00
		0.20	0.88	0.55	0.93
AYT-2	Uniform	0.60	0.91	0.65	0.95
		1.00	1.00	1.00	1.00
		0.20	0.87	0.50	0.92
	Negative Exponential	0.60	0.91	0.68	0.95
		1.00	1.00	1.00	1.00
		0.20	0.80	0.60	0.87
	3-categories QTL	0.60	0.75	0.50	0.83
		1.00	1.00	1.00	1.00
		0.20	0.75	0.50	0.83
EYT	Uniform	0.60	0.80	0.60	0.87
		1.00	1.00	1.00	1.00
		0.20	0.80	0.60	0.87
	Negative Exponential	0.60	0.75	0.50	0.83
		1.00	1.00	1.00	1.00

APPENDIX A33

ANALYSIS OF VARIANCE FOR SIMULATED GENOTYPIC VALUES FOR YIELD DUE TO STAGE OF DEVELOPMENT, FOUNDER SETS, GENETIC ARCHITECTURES AND HERITABILITY OF LINES FOR SIMULATED EXPERIMENT 4 IN WHICH CROSSES CONSISTS OF RELATIVE PREDICTIONS FROM LINES SELECTED FROM AYTS AND EYT IN EXPERIMENT 3, PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD.

Source	**Sum of Squares**	**Mean Square**	**F Ratio**	**Prob > F**
Model	74973.1	937.1	946.0	<.0001*
Error	9281.2	0.9		
C. Total	84254.4			
Source		**S.Squares**	**F Ratio**	**Prob > F**
Founder_Set		101.8	51.3	<.0001*
Stage		2808.2	1417.3	<.0001*
Founder_Set*Stage		2.9	0.7	0.5630
Genetic_Architecture		6340.8	3200.4	<.0001*
Founder_Set*Genetic_Architecture		310.4	78.3	<.0001*
Stage*Genetic_Architecture		343.0	86.5	<.0001*
Founder_Set*Stage*Genetic_Architecture		7.4	0.9	0.4803
Heritability		10336.7	5217.2	<.0001*
Founder_Set*Heritability		504.1	127.2	<.0001*
Stage*Heritability		264.6	66.7	<.0001*
Founder_Set*Stage*Heritability		8.2	1.0	0.4038
Genetic_Architecture*Heritability		341.9	86.2	<.0001*
Founder_Set*Genetic_Architecture*Heritability		863.1	108.9	<.0001*
Stage*Genetic_Architecture*Heritability		145.7	18.3	<.0001*
Founder_Set*Stage*Genetic_Archit.*Heritability		14.9	0.9	0.5183

APPENDIX A34

SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FOR EXPERIMENT 4 (GA SET).

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield				Lodging				SCN				PRR			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	65.1	65.5	66.0	65.8	67.6	67.6	68.2	68.0	3.4	3.2	3.0	3.1	1.0	1.0	1.0	1.0	1.1	1.0	1.0	1.1
		0.6	67.6	68.5	68.6	67.9	67.6	68.5	68.6	67.9	3.5	3.5	3.1	3.0	1.0	1.0	1.0	1.0	1.1	1.1	1.2	1.3
		1.0	70.9	71.1	71.7	71.9	70.9	71.1	71.7	71.9	3.3	2.9	2.9	3.2	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.1
	uniform	0.2	62.8	62.7	62.8	62.6	64.0	64.0	64.1	63.8	3.5	3.4	3.4	3.2	1.0	1.0	1.0	1.0	1.1	1.1	1.2	1.1
		0.6	64.2	64.5	64.8	64.6	64.6	65.1	65.3	65.1	3.8	3.7	3.6	3.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	67.3	67.1	67.0	66.6	67.3	67.1	67.0	66.6	3.7	3.5	3.7	3.5	1.0	1.0	1.0	1.0	1.0	1.1	1.3	1.2
	Negative Exponential	0.2	62.1	62.0	62.0	62.0	63.2	63.1	63.0	63.0	3.2	3.3	3.2	3.2	1.0	1.0	1.0	1.0	1.0	1.1	1.1	1.1
		0.6	64.1	64.8	65.2	64.2	64.5	65.2	65.5	64.7	3.2	3.1	3.2	2.8	1.0	1.0	1.0	1.0	1.1	1.1	1.1	1.4
		1.0	65.6	65.6	66.1	65.3	65.6	65.6	66.1	65.3	3.3	3.0	3.2	3.4	1.0	1.0	1.0	1.0	1.0	1.2	1.1	1.1
AYT-2	3-categories QTL	0.2	66.3	67.1	67.5	66.5	67.3	68.1	68.6	68.2	4.0	3.4	2.8	2.8	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.2
		0.6	71.4	71.8	71.8	71.2	71.4	71.8	71.8	71.2	3.8	3.4	3.1	4.1	1.0	1.0	1.0	1.0	1.2	1.2	1.4	1.4
		1.0	72.8	72.8	73.5	73.7	72.8	72.8	73.5	73.7	2.9	2.4	2.8	3.2	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	0.2	63.5	63.6	63.6	63.6	64.3	64.4	64.3	64.3	3.7	3.8	3.5	3.1	1.0	1.0	1.0	1.0	1.0	1.0	1.4	1.0
		0.6	65.2	65.6	65.7	65.7	65.5	66.0	66.1	66.0	4.0	3.8	3.8	3.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	68.2	68.2	67.8	67.7	68.2	68.2	67.8	67.7	4.2	3.8	3.6	3.4	1.0	1.0	1.0	1.0	1.0	1.0	1.6	1.0
	Negative Exponential	0.2	62.7	62.6	62.8	62.6	63.4	63.5	63.7	63.4	3.0	3.0	2.7	3.8	1.0	1.0	1.0	1.0	1.0	1.2	1.0	1.0
		0.6	65.5	66.0	66.5	65.8	66.0	66.5	66.8	66.0	2.9	3.0	3.2	3.4	1.0	1.0	1.0	1.0	1.2	1.0	1.4	1.4
		1.0	66.7	67.0	67.8	67.2	66.7	67.0	67.8	67.2	3.1	3.7	3.4	3.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
EYT	3-categories QTL	0.2	66.7	67.6	68.6	67.3	68.2	68.1	69.5	67.9	4.7	3.5	2.5	2.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	72.6	72.9	73.1	76.1	72.6	72.9	73.1	76.1	3.5	2.5	2.7	3.7	1.0	1.0	1.0	1.0	1.0	2.0	1.7	1.0
		1.0	73.8	73.2	73.9	74.1	73.8	73.2	73.9	74.1	3.0	2.3	2.7	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	0.2	63.9	64.1	64.2	64.0	65.2	64.8	64.9	64.5	3.7	4.0	3.0	3.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	66.1	66.3	66.0	66.2	66.7	66.8	66.0	66.8	4.0	3.7	4.0	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	68.6	69.3	68.1	68.4	68.6	69.3	68.1	68.4	4.3	4.0	4.0	3.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	0.2	62.7	63.0	62.8	62.5	63.5	63.7	63.3	63.3	3.0	3.0	2.0	3.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	66.2	66.7	67.3	65.9	67.1	67.2	67.5	66.7	3.0	3.0	3.3	2.5	1.0	1.0	1.0	1.0	1.0	1.0	1.7	1.0
		1.0	67.2	67.7	68.7	68.0	67.2	67.7	68.7	68.0	2.0	3.7	3.3	4.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

APPENDIX A35

SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FOR EXPERIMENT 4 (GB SET).

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield				Lodging				SCN				PRR			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	**0.2**	65.6	65.4	65.5	65.1	67.7	67.4	67.6	67.3	2.7	2.5	2.5	2.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	66.2	66.8	67.8	67.9	67.2	67.7	68.7	69.0	2.8	3.5	3.6	3.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.1
		1.0	72.6	72.1	71.8	70.8	72.6	72.1	71.8	70.8	2.2	2.1	2.3	2.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	**0.2**	63.4	63.2	63.4	63.0	64.5	64.3	64.4	64.1	2.4	2.8	2.9	2.8	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	66.4	66.1	66.4	65.9	66.9	66.6	66.9	66.4	2.7	2.9	2.6	3.2	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	68.6	68.3	68.5	68.1	68.6	68.3	68.5	68.1	3.1	3.3	3.6	3.8	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	**0.2**	62.3	62.4	62.3	62.1	63.2	63.3	63.4	63.1	3.0	2.9	3.0	3.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	66.1	65.7	66.0	65.4	66.5	66.1	66.4	66.0	2.5	2.6	2.8	3.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	67.1	67.1	67.2	66.9	67.1	67.1	67.2	66.9	2.7	3.1	3.4	3.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
AYT-2	3-categories QTL	**0.2**	66.8	66.8	66.3	66.3	68.0	67.7	67.2	67.2	2.6	3.1	2.9	2.2	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	68.1	68.2	69.6	69.6	68.5	68.6	70.3	70.2	2.7	3.6	3.8	3.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	74.6	75.3	74.4	73.7	74.6	75.3	74.4	73.7	2.3	2.0	2.9	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	**0.2**	64.5	64.6	64.4	64.0	65.1	65.1	65.1	64.7	2.0	2.7	2.8	2.8	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	67.3	66.9	67.2	66.7	67.8	67.2	67.6	67.1	2.9	3.5	2.9	3.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	69.7	69.4	69.7	69.4	69.7	69.4	69.7	69.4	3.2	2.8	4.0	4.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	**0.2**	63.1	62.9	63.1	62.7	63.6	63.5	63.8	63.3	3.1	2.8	3.3	3.2	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	67.2	67.2	66.9	66.6	67.5	67.4	67.3	67.0	2.3	2.0	2.6	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	68.6	68.9	68.4	68.1	68.6	68.9	68.4	68.1	3.1	3.1	3.1	3.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
EYT	3-categories QTL	**0.2**	67.2	67.1	67.1	66.8	68.5	68.7	68.2	67.9	2.0	3.7	3.0	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	69.3	69.0	70.2	70.9	69.8	70.8	70.8	71.1	2.5	4.5	4.0	3.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	76.0	77.3	76.3	75.0	76.0	77.3	76.3	75.0	2.3	2.0	4.0	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	**0.2**	64.6	65.1	64.9	64.6	65.2	65.5	65.9	65.4	2.0	3.0	3.3	3.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	67.7	67.5	67.4	67.3	68.0	67.5	68.0	67.8	3.3	5.0	3.7	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	70.3	69.7	70.3	69.9	70.3	69.7	70.3	69.9	2.3	3.0	4.0	4.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	**0.2**	62.9	63.4	63.4	63.0	64.8	63.5	63.9	64.1	2.0	4.0	2.3	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	67.4	67.8	67.6	67.4	67.8	68.2	68.4	67.9	2.0	2.0	2.5	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	69.1	69.4	68.8	68.9	69.1	69.4	68.8	68.9	3.0	2.3	2.7	5.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

APPENDIX A36

SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FOR EXPERIMENT 4 (GC SET).

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield				Lodging				SCN				PRR			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	65.1	64.7	64.8	65.2	67.4	66.7	66.9	66.9	2.4	2.9	3.2	3.4	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	66.2	66.9	68.1	67.1	67.3	67.9	69.0	68.1	2.6	2.8	2.7	2.7	1.0	1.0	1.0	1.0	1.0	1.1	1.2	1.3
		1.0	71.7	71.6	71.5	70.0	71.7	71.6	71.5	70.0	2.5	2.7	3.1	3.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	0.2	62.8	63.0	63.2	62.9	64.1	64.0	64.2	63.8	2.6	2.6	2.5	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	65.3	65.0	65.2	64.5	65.8	65.5	65.7	64.9	2.4	2.8	2.9	3.2	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	66.9	66.8	67.5	66.8	66.9	66.8	67.5	66.8	2.6	3.1	3.3	3.8	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	0.2	61.9	62.1	62.1	62.0	63.0	63.1	63.3	63.2	2.7	3.1	2.9	2.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	66.6	66.5	66.6	66.5	67.2	67.0	67.2	67.0	2.7	3.0	3.2	3.4	1.0	1.0	1.0	1.0	1.0	1.0	1.1	1.0
		1.0	67.9	68.3	68.4	68.1	67.9	68.3	68.4	68.1	2.3	2.9	2.9	3.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
AYT-2	3-categories QTL	0.2	66.1	66.2	65.7	66.8	67.5	66.7	66.8	67.7	2.7	2.7	3.6	2.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	68.0	68.4	70.3	69.5	68.9	68.8	70.3	70.0	2.6	2.7	2.3	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.4	1.4
		1.0	75.3	75.4	74.9	73.1	75.3	75.4	74.9	73.1	2.6	2.2	3.0	2.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	0.2	63.5	63.8	64.0	63.8	64.3	64.4	65.0	64.4	2.8	2.9	2.5	2.8	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	65.7	66.4	66.2	65.5	66.3	66.7	66.4	65.8	2.7	2.9	2.8	3.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.2
		1.0	67.9	68.3	68.7	68.1	67.9	68.3	68.7	68.1	2.4	3.1	3.4	3.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	0.2	62.5	62.8	62.7	62.8	63.3	63.5	63.4	63.6	2.6	2.5	3.2	2.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	67.4	67.4	67.4	67.4	67.7	67.7	67.9	67.7	3.1	3.3	3.1	3.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	69.8	70.0	69.9	69.2	69.8	70.0	69.9	69.2	2.0	2.3	2.8	4.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
EYT	3-categories QTL	0.2	66.8	66.2	66.1	67.2	66.9	67.9	67.3	68.1	2.0	3.3	2.5	3.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	68.5	68.9	71.3	70.9	69.9	69.9	70.8	71.1	2.5	2.5	1.7	2.3	1.0	1.0	1.0	1.0	1.0	1.0	1.7	1.7
		1.0	77.1	76.2	76.2	75.9	77.1	76.2	76.2	75.9	3.7	2.0	3.3	4.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	0.2	64.2	64.0	64.6	64.3	65.0	65.0	64.8	64.9	3.5	2.0	2.0	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	66.3	67.1	66.5	66.0	66.9	67.3	66.8	66.8	3.0	2.7	2.3	4.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	68.4	68.9	69.1	68.7	68.4	68.9	69.1	68.7	1.5	3.3	4.3	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	0.2	62.7	63.4	62.7	63.2	63.2	63.3	63.9	63.8	2.5	2.5	2.7	2.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	67.9	67.8	67.7	67.8	68.7	68.3	68.5	68.2	3.0	4.3	3.3	2.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	71.0	71.0	70.5	70.1	71.0	71.0	70.5	70.1	2.3	2.3	2.3	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

APPENDIX A37

LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 4 (GA SET), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES.

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	65.1 (±0.12)	65.5 (±0.11)	66.0 (±0.11)	65.8 (±0.11)	67.6 (±0.11)	67.6 (±0.10)	68.2 (±0.11)	68.0 (±0.11)
		0.6	67.6	68.5	68.6	67.9	67.6	68.5	68.6	67.9
		1.0	70.9	71.1	71.7	71.9	70.9	71.1	71.7	71.9
	uniform	0.2	62.8	62.7	62.8	62.6	64.0	64.0	64.1	63.8
		0.6	64.2	64.5	64.8	64.6	64.6	65.1	65.3	65.1
		1.0	67.3	67.1	67.0	66.6	67.3	67.1	67.0	66.6
	Negative Exponential	0.2	62.1	62.0	62.0	62.0	63.2	63.1	63.0	63.0
		0.6	64.1	64.8	65.2	64.2	64.5	65.2	65.5	64.7
		1.0	65.6	65.6	66.1	65.3	65.6	65.6	66.1	65.3
AYT-2	3-categories QTL	0.2	66.3 (±0.21)	67.1 (±0.22)	67.5 (±0.21)	66.5 (±0.25)	67.3 (±0.16)	68.1 (±0.21)	68.6 (±0.21)	68.2 (±0.24)
		0.6	71.4	71.8	71.8	71.2	71.4	71.8	71.8	71.2
		1.0	72.8	72.8	73.5	73.7	72.8	72.8	73.5	73.7
	uniform	0.2	63.5	63.6	63.6	63.6	64.3	64.4	64.3	64.3
		0.6	65.2	65.6	65.7	65.7	65.5	66.0	66.1	66.0
		1.0	68.2	68.2	67.8	67.7	68.2	68.2	67.8	67.7
	Negative Exponential	0.2	62.7	62.6	62.8	62.6	63.4	63.5	63.7	63.4
		0.6	65.5	66.0	66.5	65.8	66.0	66.5	66.8	66.0
		1.0	66.7	67.0	67.8	67.2	66.7	67.0	67.8	67.2
EYT	3-categories QTL	0.2	66.7 (±0.42)	67.6 (±0.43)	68.6 (±0.49)	67.3 (±0.22)	68.2 (±0.25)	68.1 (±0.34)	69.5 (±0.31)	67.9 (±0.18)
		0.6	72.6	72.9	73.1	76.1	72.6	72.9	73.1	76.1
		1.0	73.8	73.2	73.9	74.1	73.8	73.2	73.9	74.1
	uniform	0.2	63.9	64.1	64.2	64.0	65.2	64.8	64.9	64.5
		0.6	66.1	66.3	66.0	66.2	66.7	66.8	66.0	66.8
		1.0	68.6	69.3	68.1	68.4	68.6	69.3	68.1	68.4
	Negative Exponential	0.2	62.7	63.0	62.8	62.5	63.5	63.7	63.3	63.3
		0.6	66.2	66.7	67.3	65.9	67.1	67.2	67.5	66.7
		1.0	67.2	67.7	68.7	68.0	67.2	67.7	68.7	68.0

APPENDIX A38

LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 4 (GB SET), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES.

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield			
			1.8-2.4	**2.4-3.0**	**3.0-3.6**	**3.6-4.2**	**1.8-2.4**	**2.4-3.0**	**3.0-3.6**	**3.6-4.2**
AYT-1	**3-categories QTL**	**0.2**	65.6 (±0.10)	65.4 (±0.11)	65.5 (±0.10)	65.1 (±0.10)	67.7 (±0.09)	67.4 (±0.10)	67.6 (±0.09)	67.3 (±0.10)
		0.6	66.2	66.8	67.8	67.9	67.2	67.7	68.7	69.0
		1.0	72.6	72.1	71.8	70.8	72.6	72.1	71.8	70.8
	uniform	**0.2**	63.4	63.2	63.4	63.0	64.5	64.3	64.4	64.1
		0.6	66.4	66.1	66.4	65.9	66.9	66.6	66.9	66.4
		1.0	68.6	68.3	68.5	68.1	68.6	68.3	68.5	68.1
	Negative Exponential	**0.2**	62.3	62.4	62.3	62.1	63.2	63.3	63.4	63.1
		0.6	66.1	65.7	66.0	65.4	66.5	66.1	66.4	66.0
		1.0	67.1	67.1	67.2	66.9	67.1	67.1	67.2	66.9
AYT-2	**3-categories QTL**	**0.2**	66.8 (±0.19)	66.8 (±0.24)	66.3 (±0.21)	66.3 (±0.21)	68.0 (±0.19)	67.7 (±0.21)	67.2 (±0.19)	67.2 (±0.19)
		0.6	68.1	68.2	69.6	69.6	68.5	68.6	70.3	70.2
		1.0	74.6	75.3	74.4	73.7	74.6	75.3	74.4	73.7
	uniform	**0.2**	64.5	64.6	64.4	64.0	65.1	65.1	65.1	64.7
		0.6	67.3	66.9	67.2	66.7	67.8	67.2	67.6	67.1
		1.0	69.7	69.4	69.7	69.4	69.7	69.4	69.7	69.4
	Negative Exponential	**0.2**	63.1	62.9	63.1	62.7	63.6	63.5	63.8	63.3
		0.6	67.2	67.2	66.9	66.6	67.5	67.4	67.3	67.0
		1.0	68.6	68.9	68.4	68.1	68.6	68.9	68.4	68.1
EYT	**3-categories QTL**	**0.2**	67.2 (±0.32)	67.1 (±0.32)	67.1 (±0.39)	66.8 (±0.52)	68.5 (±0.25)	68.7 (±0.40)	68.2 (±0.41)	67.9 (±0.49)
		0.6	69.3	69.0	70.2	70.9	69.8	70.8	70.8	71.1
		1.0	76.0	77.3	76.3	75.0	76.0	77.3	76.3	75.0
	uniform	**0.2**	64.6	65.1	64.9	64.6	65.2	65.5	65.9	65.4
		0.6	67.7	67.5	67.4	67.3	68.0	67.5	68.0	67.8
		1.0	70.3	69.7	70.3	69.9	70.3	69.7	70.3	69.9
	Negative Exponential	**0.2**	62.9	63.4	63.4	63.0	64.8	63.5	63.9	64.1
		0.6	67.4	67.8	67.6	67.4	67.8	68.2	68.4	67.9
		1.0	69.1	69.4	68.8	68.9	69.1	69.4	68.8	68.9

APPENDIX A39

LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 4 (GC SET), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES.

Stage	Yield Genetic Architecture		Genotypic yield				Phenotypic yield			
		H^2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	65.1 (±0.11)	64.7 (±0.12)	64.8 (±0.12)	65.2 (±0.11)	67.4 (±0.10)	66.7 (±0.12)	66.9 (±0.11)	66.9 (±0.10)
		0.6	66.2	66.9	68.1	67.1	67.3	67.9	69.0	68.1
		1.0	71.7	71.6	71.5	70.0	71.7	71.6	71.5	70.0
	uniform	0.2	62.8	63.0	63.2	62.9	64.1	64.0	64.2	63.8
		0.6	65.3	65.0	65.2	64.5	65.8	65.5	65.7	64.9
		1.0	66.9	66.8	67.5	66.8	66.9	66.8	67.5	66.8
	Negative Exponential	0.2	61.9	62.1	62.1	62.0	63.0	63.1	63.3	63.2
		0.6	66.6	66.5	66.6	66.5	67.2	67.0	67.2	67.0
		1.0	67.9	68.3	68.4	68.1	67.9	68.3	68.4	68.1
AYT-2	3-categories QTL	0.2	66.1 (±0.24)	66.2 (±0.19)	65.7 (±0.21)	66.8 (±0.20)	67.5 (±0.23)	66.7 (±0.17)	66.8 (±0.18)	67.7 (±0.18)
		0.6	68.0	68.4	70.3	69.5	68.9	68.8	70.3	70.0
		1.0	75.3	75.4	74.9	73.1	75.3	75.4	74.9	73.1
	uniform	0.2	63.5	63.8	64.0	63.8	64.3	64.4	65.0	64.4
		0.6	65.7	66.4	66.2	65.5	66.3	66.7	66.4	65.8
		1.0	67.9	68.3	68.7	68.1	67.9	68.3	68.7	68.1
	Negative Exponential	0.2	62.5	62.8	62.7	62.8	63.3	63.5	63.4	63.6
		0.6	67.4	67.4	67.4	67.4	67.7	67.7	67.9	67.7
		1.0	69.8	70.0	69.9	69.2	69.8	70.0	69.9	69.2
EYT	3-categories QTL	0.2	66.8 (±0.52)	66.2 (±0.31)	66.1 (±0.30)	67.2 (±0.11)	66.9 (±0.50)	67.9 (±0.29)	67.3 (±0.32)	68.1 (±0.29)
		0.6	68.5	68.9	71.3	70.9	69.9	69.9	70.8	71.1
		1.0	77.1	76.2	76.2	75.9	77.1	76.2	76.2	75.9
	uniform	0.2	64.2	64.0	64.6	64.3	65.0	65.0	64.8	64.9
		0.6	66.3	67.1	66.5	66.0	66.9	67.3	66.8	66.8
		1.0	68.4	68.9	69.1	68.7	68.4	68.9	69.1	68.7
	Negative Exponential	0.2	62.7	63.4	62.7	63.2	63.2	63.3	63.9	63.8
		0.6	67.9	67.8	67.7	67.8	68.7	68.3	68.5	68.2
		1.0	71.0	71.0	70.5	70.1	71.0	71.0	70.5	70.1

APPENDIX A40

DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY METRICS FOR GA FOUNDING SETS OF CROSSES USED TO INITIATE VARIETY DEVELOPMENT FOR EACH HERITABILITY AND GENETIC ARCHITECTURE COMBINATION AND STAGE OF DEVELOPMENT FOR EXPERIMENT 4. THE THREE SETS OF CROSSES WERE SELECTED BASED ON PREDICTED PROGENY VALUES, WHERE THE TRAINING SET CONSISTED OF RELATED LINES SELECTED FROM AYTS AND EYT IN EXPERIMENT 3, PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD.

Stage	Yield Genetic Architecture	H^2	Accuracy	Sensitivity	Specificity
PYT	3-categories QTL	0.20	0.87	0.37	0.93
		0.60	0.93	0.64	0.96
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.87	0.37	0.93
		0.60	0.94	0.71	0.97
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.87	0.34	0.93
		0.60	0.94	0.70	0.97
		1.00	1.00	1.00	1.00
AYT-1	3-categories QTL	0.20	0.80	0.48	0.88
		0.60	0.85	0.60	0.90
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.84	0.58	0.90
		0.60	0.88	0.70	0.93
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.82	0.54	0.89
		0.60	0.89	0.73	0.93
		1.00	1.00	1.00	1.00
AYT-2	3-categories QTL	0.20	0.85	0.45	0.92
		0.60	0.90	0.63	0.94
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.86	0.48	0.92
		0.60	0.91	0.65	0.95
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.87	0.50	0.92
		0.60	0.92	0.70	0.95
		1.00	1.00	1.00	1.00
EYT	3-categories QTL	0.20	0.75	0.50	0.83
		0.60	0.75	0.50	0.83
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.85	0.70	0.90
		0.60	0.80	0.60	0.87
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.60	0.20	0.73
		0.60	0.80	0.60	0.87
		1.00	1.00	1.00	1.00

APPENDIX A41

DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY METRICS FOR GB FOUNDING SETS OF CROSSES USED TO INITIATE VARIETY DEVELOPMENT FOR EACH HERITABILITY AND GENETIC ARCHITECTURE COMBINATION AND STAGE OF DEVELOPMENT FOR EXPERIMENT 4. THE THREE SETS OF CROSSES WERE SELECTED BASED ON PREDICTED PROGENY VALUES, WHERE THE TRAINING SET CONSISTED OF RELATED LINES SELECTED FROM AYTS AND EYT IN EXPERIMENT 3, PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD.

Stage	Yield Genetic Architecture	H^2	Accuracy	Sensitivity	Specificity
PYT	3-categories QTL	0.20	0.87	0.34	0.93
		0.60	0.93	0.66	0.96
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.88	0.40	0.93
		0.60	0.93	0.63	0.96
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.88	0.38	0.93
		0.60	0.93	0.66	0.96
		1.00	1.00	1.00	1.00
AYT-1	3-categories QTL	0.20	0.79	0.47	0.87
		0.60	0.85	0.62	0.91
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.83	0.57	0.90
		0.60	0.88	0.69	0.93
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.82	0.53	0.89
		0.60	0.89	0.71	0.93
		1.00	1.00	1.00	1.00
AYT-2	3-categories QTL	0.20	0.87	0.50	0.92
		0.60	0.91	0.65	0.95
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.87	0.50	0.92
		0.60	0.89	0.58	0.93
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.88	0.55	0.93
		0.60	0.93	0.73	0.96
		1.00	1.00	1.00	1.00
EYT	3-categories QTL	0.20	0.70	0.40	0.80
		0.60	0.90	0.80	0.93
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.80	0.60	0.87
		0.60	0.80	0.60	0.87
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.70	0.40	0.80
		0.60	0.90	0.80	0.93
		1.00	1.00	1.00	1.00

APPENDIX A42

DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY METRICS FOR GC FOUNDING SETS OF CROSSES USED TO INITIATE VARIETY DEVELOPMENT FOR EACH HERITABILITY AND GENETIC ARCHITECTURE COMBINATION AND STAGE OF DEVELOPMENT FOR EXPERIMENT 4. THE THREE SETS OF CROSSES WERE SELECTED BASED ON PREDICTED PROGENY VALUES, WHERE THE TRAINING SET CONSISTED OF RELATED LINES SELECTED FROM AYTS AND EYT IN EXPERIMENT 3, PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD.

Stage	Yield Genetic Architecture	H^2	Accuracy	Sensitivity	Specificity
PYT	3-categories QTL	0.20	0.87	0.35	0.93
		0.60	0.93	0.66	0.96
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.89	0.43	0.94
		0.60	0.94	0.70	0.97
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.87	0.36	0.93
		0.60	0.92	0.61	0.96
		1.00	1.00	1.00	1.00
AYT-1	3-categories QTL	0.20	0.80	0.50	0.88
		0.60	0.87	0.67	0.92
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.84	0.58	0.90
		0.60	0.88	0.70	0.93
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.80	0.49	0.88
		0.60	0.86	0.63	0.91
		1.00	1.00	1.00	1.00
AYT-2	3-categories QTL	0.20	0.87	0.50	0.92
		0.60	0.90	0.63	0.94
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.84	0.40	0.91
		0.60	0.89	0.58	0.93
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.85	0.43	0.91
		0.60	0.91	0.65	0.95
		1.00	1.00	1.00	1.00
EYT	3-categories QTL	0.20	0.70	0.40	0.80
		0.60	0.85	0.70	0.90
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.85	0.70	0.90
		0.60	0.90	0.80	0.93
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.80	0.60	0.87
		0.60	0.80	0.60	0.87
		1.00	1.00	1.00	1.00

APPENDIX A43

ANALYSIS OF VARIANCE OF AVERAGE GENOTYPIC VALUES FOR YIELD SHOWING SIGNIFICANCE FOR THE INTERACTION BETWEEN FOUNDER SET, GENETIC ARCHITECTURE AND HERITABILITY, AS WELL STAGE OF DEVELOPMENT, GENETIC ARCHITECTURE AND HERITABILITY. EXPERIMENT 5 CONSISTS OF RELATIVE PREDICTIONS FROM LINES SELECTED FROM AYTS AND EYT IN EXPERIMENT 3, PLUS MAS FOR SCN AND PRR TRAITS PLUS GENOTYPIC SELECTION FOR YIELD USING GEBVS FOR PYT LINES.

Source	**Sum of Squares**	**Mean Square**	**F Ratio**	**Prob > F**
Model	72512.4	906.4	969.9	<.0001*
Error	8755.1	0.9		
C. Total	81267.6			
Source		**S.Squares**	**F Ratio**	**Prob > F**
Founder_Set		98.2	52.5	<.0001*
Stage		2665.0	1425.9	<.0001*
Founder_Set*Stage		4.3	1.1	0.3269
Genetic_Architecture		6386.9	3417.3	<.0001*
Founder_Set*Genetic_Architecture		314.7	84.2	<.0001*
Stage*Genetic_Architecture		275.8	73.7	<.0001*
Founder_Set*Stage*Genetic_Architecture		7.7	1.0	0.4098
Heritability		9952.1	5324.9	<.0001*
Founder_Set*Heritability		508.9	136.1	<.0001*
Stage*Heritability		349.1	93.4	<.0001*
Founder_Set*Stage*Heritability		7.6	1.0	0.4139
Genetic_Architecture*Heritability		339.5	90.8	<.0001*
Founder_Set*Genetic_Architecture*Heritability		824.3	110.2	<.0001*
Stage*Genetic_Architecture*Heritability		196.4	26.2	<.0001*
Founder_Set*Stage*Genetic_Archit.* Heritability		20.2	1.3	0.1557

APPENDIX A44

SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FOR EXPERIMENT 5 (GA SET).

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield				Lodging				SCN				PRR			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	65.5	65.8	66.3	66.1	67.4	67.6	68.2	68.0	3.5	3.2	3.0	3.1	1.0	1.0	1.0	1.0	1.1	1.0	1.0	1.1
		0.6	67.5	68.5	68.6	67.8	67.5	68.5	68.6	67.8	3.5	3.4	3.3	3.2	1.0	1.0	1.0	1.0	1.1	1.1	1.2	1.3
		1.0	71.0	71.0	71.7	71.8	71.0	71.0	71.7	71.8	3.3	2.9	2.8	3.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	0.2	62.8	62.9	63.0	62.7	64.2	64.2	64.3	64.0	3.6	3.5	3.5	3.4	1.0	1.0	1.0	1.0	1.1	1.2	1.2	1.1
		0.6	64.3	64.5	64.8	64.6	64.6	65.1	65.3	65.1	3.8	3.7	3.5	3.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	67.3	67.1	67.0	66.6	67.3	67.1	67.0	66.6	3.7	3.5	3.7	3.5	1.0	1.0	1.0	1.0	1.0	1.1	1.2	1.2
	Negative Exponential	0.2	62.1	62.1	62.1	62.0	63.0	63.2	63.1	63.1	3.4	3.6	3.2	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.1	1.1
		0.6	64.1	64.8	65.2	64.2	64.5	65.2	65.6	64.7	3.2	3.1	3.2	2.9	1.0	1.0	1.0	1.0	1.1	1.2	1.1	1.3
		1.0	65.6	65.6	66.0	65.3	65.6	65.6	66.0	65.3	3.3	3.1	3.0	3.4	1.0	1.0	1.0	1.0	1.0	1.2	1.1	1.1
AYT-2	3-categories QTL	0.2	66.6	67.0	67.2	66.9	67.5	68.2	68.4	67.7	3.8	3.2	3.1	2.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.2
		0.6	71.4	71.8	71.8	71.2	71.4	71.8	71.8	71.2	3.8	3.4	3.3	4.1	1.0	1.0	1.0	1.0	1.2	1.2	1.2	1.4
		1.0	72.8	72.9	73.3	73.5	72.8	72.9	73.3	73.5	2.9	2.4	2.8	3.2	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	0.2	63.7	63.8	63.7	63.6	64.2	64.4	64.4	64.2	3.6	3.5	3.2	3.0	1.0	1.0	1.0	1.0	1.0	1.2	1.0	1.0
		0.6	65.2	65.6	65.8	65.7	65.5	66.0	66.0	66.0	4.0	3.8	3.9	3.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	68.2	68.2	67.8	67.7	68.2	68.2	67.8	67.7	4.2	3.8	3.6	3.4	1.0	1.0	1.0	1.0	1.0	1.0	1.6	1.0
	Negative Exponential	0.2	62.7	62.6	62.8	62.7	63.4	63.4	63.6	63.4	2.9	3.0	2.7	3.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.2
		0.6	65.5	66.0	66.5	65.8	66.0	66.5	66.8	66.0	2.9	3.0	3.2	3.4	1.0	1.0	1.0	1.0	1.2	1.0	1.4	1.4
		1.0	66.7	67.0	67.8	67.2	66.7	67.0	67.8	67.2	3.1	3.7	3.4	3.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
EYT	3-categories QTL	0.2	67.3	67.7	68.6	67.3	68.4	68.0	69.5	67.9	4.5	3.3	2.5	2.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	72.6	72.9	73.1	76.1	72.6	72.9	73.1	76.1	3.5	2.7	3.5	3.7	1.0	1.0	1.0	1.0	1.0	1.7	1.0	1.7
		1.0	73.8	73.1	73.9	73.9	73.8	73.1	73.9	73.9	3.0	2.3	2.7	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	0.2	63.9	64.4	64.3	64.2	65.2	64.8	64.6	64.6	3.7	3.3	3.0	2.5	1.0	1.0	1.0	1.0	1.0	1.7	1.0	1.0
		0.6	66.1	66.3	66.0	66.2	66.7	66.8	66.2	66.8	4.0	3.7	4.0	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	68.6	69.3	68.1	68.4	68.6	69.3	68.1	68.4	4.3	4.0	4.0	3.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	0.2	62.6	63.2	62.9	62.8	63.5	63.8	63.6	63.6	3.0	2.0	2.0	2.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	2.0
		0.6	66.2	66.7	67.3	65.9	67.1	67.2	67.5	66.7	3.0	3.0	3.3	2.5	1.0	1.0	1.0	1.0	1.0	1.0	1.7	1.0
		1.0	67.2	67.7	68.7	68.0	67.2	67.7	68.7	68.0	2.0	3.7	3.3	4.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

APPENDIX A45

SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FOR EXPERIMENT 5 (GB SET).

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield				Lodging				SCN				PRR			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	**0.2**	66.1	65.9	65.8	65.4	67.7	67.5	67.8	67.3	2.7	2.7	2.6	2.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	66.1	66.7	67.9	67.9	67.0	67.6	68.5	68.9	2.8	3.6	3.4	3.5	1.0	1.0	1.0	1.0	1.1	1.0	1.0	1.1
		1.0	72.6	72.1	71.8	70.8	72.6	72.1	71.8	70.8	2.2	2.1	2.3	2.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	**0.2**	63.4	63.4	63.4	63.2	64.8	64.6	64.8	64.3	2.4	2.4	2.9	2.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	66.4	66.1	66.3	65.9	66.8	66.6	66.8	66.4	2.7	2.9	2.7	3.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	68.6	68.3	68.5	68.1	68.6	68.3	68.5	68.1	3.1	3.3	3.5	3.8	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	**0.2**	62.5	62.4	62.5	62.2	63.3	63.4	63.5	63.2	3.0	2.8	2.6	2.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	66.0	65.7	66.0	65.5	66.4	66.1	66.4	66.0	2.5	2.8	2.8	3.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	67.1	67.1	67.1	66.9	67.1	67.1	67.1	66.9	2.8	3.1	3.4	4.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
AYT-2	3-categories QTL	**0.2**	66.8	66.8	66.2	66.2	68.2	67.6	67.6	67.4	2.9	3.4	3.1	2.4	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	68.0	68.1	69.6	69.8	68.4	68.6	70.3	70.2	3.0	3.4	3.6	3.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	74.6	75.3	74.4	73.7	74.6	75.3	74.4	73.7	2.3	2.0	2.9	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	**0.2**	64.4	64.6	64.6	64.2	65.0	65.1	65.2	64.8	2.2	2.8	2.9	2.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	67.3	66.9	67.2	66.7	67.8	67.2	67.5	67.0	2.9	3.4	3.0	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	69.7	69.4	69.7	69.4	69.7	69.4	69.7	69.4	3.2	2.8	4.0	4.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	**0.2**	63.0	62.9	63.2	62.7	63.5	63.5	63.8	63.2	3.4	3.3	3.4	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	67.1	67.2	66.9	66.6	67.4	67.4	67.3	67.0	2.5	2.0	2.6	3.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.2
		1.0	68.6	68.9	68.3	68.1	68.6	68.9	68.3	68.1	3.1	3.1	3.5	3.8	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
EYT	3-categories QTL	**0.2**	67.4	66.8	66.9	66.3	68.4	68.5	68.0	67.2	2.3	3.3	3.3	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	69.3	69.0	70.2	70.9	69.8	70.8	70.8	71.1	2.5	4.5	4.0	3.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	76.0	77.3	75.8	75.5	76.0	77.3	75.8	75.5	2.3	2.0	4.0	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	**0.2**	64.6	65.3	64.9	64.8	65.2	65.8	65.9	65.3	2.0	3.3	3.3	3.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	67.7	67.5	67.4	67.3	68.0	67.5	68.0	67.8	3.3	5.0	3.7	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	70.3	69.7	70.3	69.9	70.3	69.7	70.3	69.9	2.3	3.0	4.0	4.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	**0.2**	63.0	62.8	63.6	63.1	64.4	63.7	64.0	64.0	2.7	4.0	3.7	3.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	67.4	67.8	67.6	67.4	67.8	68.2	68.4	67.9	2.0	2.0	2.5	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	69.1	69.4	68.8	68.9	69.1	69.4	68.8	68.9	3.0	2.3	2.7	5.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

APPENDIX A46

SIMULATED GENOTYPIC VALUES FOR YIELD, PHENOTYPIC YIELD, LODGING, SCN AND PRR FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT FOR EXPERIMENT 5 (GC SET).

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield				Lodging				SCN				PRR			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	65.4	65.2	65.4	65.3	67.3	66.7	67.1	67.1	2.5	3.1	3.0	3.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	66.2	66.9	68.1	67.2	67.1	67.8	68.8	68.0	2.7	2.8	2.7	2.7	1.0	1.0	1.0	1.0	1.0	1.1	1.2	1.3
		1.0	71.7	71.6	71.5	70.0	71.7	71.6	71.5	70.0	2.5	2.7	3.1	3.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	0.2	62.8	63.0	63.3	63.0	64.3	64.2	64.6	64.1	2.4	2.5	2.3	2.8	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	65.3	65.0	65.2	64.5	65.8	65.5	65.7	64.9	2.5	2.7	2.9	3.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	66.9	66.8	67.5	66.8	66.9	66.8	67.5	66.8	2.6	3.1	3.3	3.8	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	0.2	62.0	62.3	62.2	62.2	63.2	63.2	63.4	63.3	2.6	2.7	2.7	2.8	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	66.5	66.5	66.6	66.4	67.2	66.9	67.1	67.0	2.8	3.1	3.4	3.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	67.8	68.3	68.3	68.0	67.8	68.3	68.3	68.0	2.4	2.9	3.1	3.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
AYT-2	3-categories QTL	0.2	66.1	66.1	65.9	66.7	67.6	67.0	67.2	67.6	2.7	3.2	3.0	2.4	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	68.0	68.4	70.3	69.5	68.8	68.8	70.3	70.0	2.7	2.7	2.3	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.4	1.4
		1.0	75.3	75.4	74.9	73.1	75.3	75.4	74.9	73.1	2.6	2.2	3.0	2.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	0.2	63.6	63.8	64.1	63.8	64.4	64.4	65.1	64.5	2.5	2.8	2.2	2.8	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	65.7	66.4	66.2	65.5	66.3	66.7	66.4	65.8	2.7	2.9	2.9	3.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.2
		1.0	67.9	68.3	68.7	68.1	67.9	68.3	68.7	68.1	2.4	3.1	3.4	3.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	0.2	62.4	63.0	62.7	62.8	63.3	63.8	63.3	63.4	2.3	2.8	2.7	1.8	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	67.3	67.5	67.3	67.2	67.7	67.7	67.9	67.6	2.9	3.0	3.6	3.8	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	69.8	70.0	69.9	69.2	69.8	70.0	69.9	69.2	2.0	2.3	2.8	4.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
EYT	3-categories QTL	0.2	66.6	66.4	65.6	67.1	67.4	67.8	67.5	67.8	2.0	3.3	2.0	2.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	68.3	69.1	71.3	70.9	70.3	69.8	70.8	71.1	1.0	3.0	1.7	2.3	1.0	1.0	1.0	1.0	1.0	1.0	1.7	1.7
		1.0	77.1	76.2	76.2	75.9	77.1	76.2	76.2	75.9	3.7	2.0	3.3	4.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	uniform	0.2	64.2	64.0	65.1	64.3	65.0	65.2	65.2	64.9	3.5	2.0	2.0	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	66.3	67.1	66.5	65.8	66.9	67.3	66.8	66.4	3.5	2.7	2.3	2.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	68.4	68.9	69.1	68.7	68.4	68.9	69.1	68.7	1.5	3.7	4.3	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Negative Exponential	0.2	62.6	63.4	62.9	63.2	63.8	63.6	63.4	63.8	1.0	2.0	2.0	2.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		0.6	67.8	67.8	67.8	67.5	68.6	68.3	68.4	68.0	4.0	4.3	3.3	4.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		1.0	71.0	71.0	70.5	70.1	71.0	71.0	70.5	70.1	2.3	2.3	2.3	3.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

APPENDIX A47

LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 5 (GA SET), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES.

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	65.5 (±0.11)	65.8 (±0.11)	66.3 (±0.11)	66.1 (±0.11)	67.4 (±0.11)	67.6 (±0.10)	68.2 (±0.11)	68.0 (±0.11)
		0.6	67.5	68.5	68.6	67.8	67.5	68.5	68.6	67.8
		1.0	71.0	71.0	71.7	71.8	71.0	71.0	71.7	71.8
	uniform	0.2	62.8	62.9	63.0	62.7	64.2	64.2	64.3	64.0
		0.6	64.3	64.5	64.8	64.6	64.6	65.1	65.3	65.1
		1.0	67.3	67.1	67.0	66.6	67.3	67.1	67.0	66.6
	Negative Exponential	0.2	62.1	62.1	62.1	62.0	63.0	63.2	63.1	63.1
		0.6	64.1	64.8	65.2	64.2	64.5	65.2	65.6	64.7
		1.0	65.6	65.6	66.0	65.3	65.6	65.6	66.0	65.3
AYT-2	3-categories QTL	0.2	66.6 (±0.20)	67.0 (±0.22)	67.2 (±0.22)	66.9 (±0.25)	67.5 (±0.16)	68.2 (±0.21)	68.4 (±0.20)	67.7 (±0.25)
		0.6	71.4	71.8	71.8	71.2	71.4	71.8	71.8	71.2
		1.0	72.8	72.9	73.3	73.5	72.8	72.9	73.3	73.5
	uniform	0.2	63.7	63.8	63.7	63.6	64.2	64.4	64.4	64.2
		0.6	65.2	65.6	65.8	65.7	65.5	66.0	66.0	66.0
		1.0	68.2	68.2	67.8	67.7	68.2	68.2	67.8	67.7
	Negative Exponential	0.2	62.7	62.6	62.8	62.7	63.4	63.4	63.6	63.4
		0.6	65.5	66.0	66.5	65.8	66.0	66.5	66.8	66.0
		1.0	66.7	67.0	67.8	67.2	66.7	67.0	67.8	67.2
EYT	3-categories QTL	0.2	67.3 (±0.46)	67.7 (±0.32)	68.6 (±0.49)	67.3 (±0.21)	68.4 (±0.30)	68.0 (±0.25)	69.5 (±0.33)	67.9 (±0.17)
		0.6	72.6	72.9	73.1	76.1	72.6	72.9	73.1	76.1
		1.0	73.8	73.1	73.9	73.9	73.8	73.1	73.9	73.9
	uniform	0.2	63.9	64.4	64.3	64.2	65.2	64.8	64.6	64.6
		0.6	66.1	66.3	66.0	66.2	66.7	66.8	66.2	66.8
		1.0	68.6	69.3	68.1	68.4	68.6	69.3	68.1	68.4
	Negative Exponential	0.2	62.6	63.2	62.9	62.8	63.5	63.8	63.6	63.6
		0.6	66.2	66.7	67.3	65.9	67.1	67.2	67.5	66.7
		1.0	67.2	67.7	68.7	68.0	67.2	67.7	68.7	68.0

APPENDIX A48

LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 5 (GB SET), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES.

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	66.1 (±0.09)	65.9 (±0.10)	65.8 (±0.10)	65.4 (±0.10)	67.7 (±0.08)	67.5 (±0.10)	67.8 (±0.09)	67.3 (±0.10)
		0.6	66.1	66.7	67.9	67.9	67.0	67.6	68.5	68.9
		1.0	72.6	72.1	71.8	70.8	72.6	72.1	71.8	70.8
	uniform	0.2	63.4	63.4	63.4	63.2	64.8	64.6	64.8	64.3
		0.6	66.4	66.1	66.3	65.9	66.8	66.6	66.8	66.4
		1.0	68.6	68.3	68.5	68.1	68.6	68.3	68.5	68.1
	Negative Exponential	0.2	62.5	62.4	62.5	62.2	63.3	63.4	63.5	63.2
		0.6	66.0	65.7	66.0	65.5	66.4	66.1	66.4	66.0
		1.0	67.1	67.1	67.1	66.9	67.1	67.1	67.1	66.9
AYT-2	3-categories QTL	0.2	66.8 (±0.20)	66.8 (±0.23)	66.2 (±0.20)	66.2 (±0.20)	68.2 (±0.18)	67.6 (±0.21)	67.6 (±0.18)	67.4 (±0.18)
		0.6	68.0	68.1	69.6	69.8	68.4	68.6	70.3	70.2
		1.0	74.6	75.3	74.4	73.7	74.6	75.3	74.4	73.7
	uniform	0.2	64.4	64.6	64.6	64.2	65.0	65.1	65.2	64.8
		0.6	67.3	66.9	67.2	66.7	67.8	67.2	67.5	67.0
		1.0	69.7	69.4	69.7	69.4	69.7	69.4	69.7	69.4
	Negative Exponential	0.2	63.0	62.9	63.2	62.7	63.5	63.5	63.8	63.2
		0.6	67.1	67.2	66.9	66.6	67.4	67.4	67.3	67.0
		1.0	68.6	68.9	68.3	68.1	68.6	68.9	68.3	68.1
EYT	3-categories QTL	0.2	67.4 (±0.25)	66.8 (±0.32)	66.9 (±0.34)	66.3 (±0.54)	68.4 (±0.27)	68.5 (±0.45)	68.0 (±0.36)	67.2 (±0.60)
		0.6	69.3	69.0	70.2	70.9	69.8	70.8	70.8	71.1
		1.0	76.0	77.3	75.8	75.5	76.0	77.3	75.8	75.5
	uniform	0.2	64.6	65.3	64.9	64.8	65.2	65.8	65.9	65.3
		0.6	67.7	67.5	67.4	67.3	68.0	67.5	68.0	67.8
		1.0	70.3	69.7	70.3	69.9	70.3	69.7	70.3	69.9
	Negative Exponential	0.2	63.0	62.8	63.6	63.1	64.4	63.7	64.0	64.0
		0.6	67.4	67.8	67.6	67.4	67.8	68.2	68.4	67.9
		1.0	69.1	69.4	68.8	68.9	69.1	69.4	68.8	68.9

APPENDIX A49

LEAST SQUARE MEANS AND STANDARD ERROR (±SE) FOR GENOTYPIC VALUES FOR YIELD AND SIMULATED PHENOTYPIC YIELD FROM EXPERIMENT 5 (GC SET), FOR ALL NINE GENETIC ARCHITECTURE AND STAGES OF DEVELOPMENT. VALUES ARE LISTED BY RM GROUP AND VARIETAL DEVELOPMENT STAGES.

Stage	Yield Genetic Architecture	H^2	Genotypic yield				Phenotypic yield			
			1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2	1.8-2.4	2.4-3.0	3.0-3.6	3.6-4.2
AYT-1	3-categories QTL	0.2	65.4 (±0.11)	65.2 (±0.12)	65.4 (±0.11)	65.3 (±0.10)	67.3 (±0.11)	66.7 (±0.11)	67.1 (±0.11)	67.1 (±0.10)
		0.6	66.2	66.9	68.1	67.2	67.1	67.8	68.8	68.0
		1.0	71.7	71.6	71.5	70.0	71.7	71.6	71.5	70.0
	uniform	0.2	62.8	63.0	63.3	63.0	64.3	64.2	64.6	64.1
		0.6	65.3	65.0	65.2	64.5	65.8	65.5	65.7	64.9
		1.0	66.9	66.8	67.5	66.8	66.9	66.8	67.5	66.8
	Negative Exponential	0.2	62.0	62.3	62.2	62.2	63.2	63.2	63.4	63.3
		0.6	66.5	66.5	66.6	66.4	67.2	66.9	67.1	67.0
		1.0	67.8	68.3	68.3	68.0	67.8	68.3	68.3	68.0
AYT-2	3-categories QTL	0.2	66.1 (±0.25)	66.1 (±0.19)	65.9 (±0.22)	66.7 (±0.20)	67.6 (±0.23)	67.0 (±0.17)	67.2 (±0.18)	67.6 (±0.19)
		0.6	68.0	68.4	70.3	69.5	68.8	68.8	70.3	70.0
		1.0	75.3	75.4	74.9	73.1	75.3	75.4	74.9	73.1
	uniform	0.2	63.6	63.8	64.1	63.8	64.4	64.4	65.1	64.5
		0.6	65.7	66.4	66.2	65.5	66.3	66.7	66.4	65.8
		1.0	67.9	68.3	68.7	68.1	67.9	68.3	68.7	68.1
	Negative Exponential	0.2	62.4	63.0	62.7	62.8	63.3	63.8	63.3	63.4
		0.6	67.3	67.5	67.3	67.2	67.7	67.7	67.9	67.6
		1.0	69.8	70.0	69.9	69.2	69.8	70.0	69.9	69.2
EYT	3-categories QTL	0.2	66.6 (±0.58)	66.4 (±0.29)	65.6 (±0.26)	67.1 (±0.15)	67.4 (±0.58)	67.8 (±0.27)	67.5 (±0.24)	67.8 (±0.36)
		0.6	68.3	69.1	71.3	70.9	70.3	69.8	70.8	71.1
		1.0	77.1	76.2	76.2	75.9	77.1	76.2	76.2	75.9
	uniform	0.2	64.2	64.0	65.1	64.3	65.0	65.2	65.2	64.9
		0.6	66.3	67.1	66.5	65.8	66.9	67.3	66.8	66.4
		1.0	68.4	68.9	69.1	68.7	68.4	68.9	69.1	68.7
	Negative Exponential	0.2	62.6	63.4	62.9	63.2	63.8	63.6	63.4	63.8
		0.6	67.8	67.8	67.8	67.5	68.6	68.3	68.4	68.0
		1.0	71.0	71.0	70.5	70.1	71.0	71.0	70.5	70.1

APPENDIX A50

AVERAGE DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY FOR GA FOUNDING SETS OF CROSSES USED TO INITIATE VARIETY DEVELOPMENT FOR NINE COMBINATIONS OF GENETIC ARCHITECTURE AND HERITABILITY AND STAGE OF DEVELOPMENT IN SOYBEAN VARIETY DEVELOPMENT FOR EXPERIMENT 5 IN WHICH CROSSES CONSISTS OF RELATIVE PREDICTIONS FROM LINES SELECTED FROM AYTS AND EYT IN EXPERIMENT 3, PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD.

Stage	Yield Genetic Architecture	H^2	Accuracy	Sensitivity	Specificity
PYT	3-categories QTL	0.20	0.91	0.53	0.95
		0.60	0.91	0.57	0.95
		1.00	0.93	0.67	0.96
	Uniform	0.20	0.91	0.53	0.95
		0.60	0.94	0.68	0.96
		1.00	0.93	0.68	0.96
	Negative Exponential	0.20	0.88	0.39	0.93
		0.60	0.93	0.64	0.96
		1.00	0.91	0.54	0.95
AYT-1	3-categories QTL	0.20	0.80	0.49	0.88
		0.60	0.86	0.64	0.91
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.81	0.51	0.88
		0.60	0.88	0.69	0.93
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.80	0.49	0.88
		0.60	0.89	0.73	0.93
		1.00	1.00	1.00	1.00
AYT-2	3-categories QTL	0.20	0.87	0.50	0.92
		0.60	0.90	0.63	0.94
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.87	0.53	0.93
		0.60	0.91	0.68	0.95
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.88	0.55	0.93
		0.60	0.92	0.70	0.95
		1.00	1.00	1.00	1.00
EYT	3-categories QTL	0.20	0.75	0.50	0.83
		0.60	0.75	0.50	0.83
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.85	0.70	0.90
		0.60	0.80	0.60	0.87
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.75	0.50	0.83
		0.60	0.80	0.60	0.87
		1.00	1.00	1.00	1.00

APPENDIX A51

AVERAGE DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY FOR GB FOUNDING SETS OF CROSSES USED TO INITIATE VARIETY DEVELOPMENT FOR NINE COMBINATIONS OF GENETIC ARCHITECTURE AND HERITABILITY AND STAGE OF DEVELOPMENT IN SOYBEAN VARIETY DEVELOPMENT FOR EXPERIMENT 5 IN WHICH CROSSES CONSISTS OF RELATIVE PREDICTIONS FROM LINES SELECTED FROM AYTS AND EYT IN EXPERIMENT 3, PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD.

Stage	Yield Genetic Architecture	H^2	Accuracy	Sensitivity	Specificity
PYT	3-categories QTL	0.20	0.85	0.27	0.92
		0.60	0.89	0.46	0.94
		1.00	0.94	0.70	0.97
	Uniform	0.20	0.86	0.29	0.92
		0.60	0.89	0.46	0.94
		1.00	0.94	0.69	0.97
	Negative Exponential	0.20	0.86	0.28	0.92
		0.60	0.89	0.47	0.94
		1.00	0.92	0.61	0.96
AYT-1	3-categories QTL	0.20	0.78	0.44	0.86
		0.60	0.87	0.68	0.92
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.81	0.51	0.88
		0.60	0.88	0.70	0.93
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.80	0.49	0.88
		0.60	0.89	0.73	0.93
		1.00	1.00	1.00	1.00
AYT-2	3-categories QTL	0.20	0.85	0.43	0.91
		0.60	0.91	0.65	0.95
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.89	0.58	0.93
		0.60	0.89	0.58	0.93
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.87	0.50	0.92
		0.60	0.93	0.73	0.96
		1.00	1.00	1.00	1.00
EYT	3-categories QTL	0.20	0.70	0.40	0.80
		0.60	0.90	0.80	0.93
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.80	0.60	0.87
		0.60	0.80	0.60	0.87
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.65	0.30	0.77
		0.60	0.90	0.80	0.93
		1.00	1.00	1.00	1.00

APPENDIX A52

AVERAGE DECISION ACCURACY, SENSITIVITY, AND SPECIFICITY FOR GC FOUNDING SETS OF CROSSES USED TO INITIATE VARIETY DEVELOPMENT FOR NINE COMBINATIONS OF GENETIC ARCHITECTURE AND HERITABILITY AND STAGE OF DEVELOPMENT IN SOYBEAN VARIETY DEVELOPMENT FOR EXPERIMENT 5 IN WHICH CROSSES CONSISTS OF RELATIVE PREDICTIONS FROM LINES SELECTED FROM AYTS AND EYT IN EXPERIMENT 3, PLUS MAS FOR SCN AND PRR TRAITS PLUS PHENOTYPIC SELECTION FOR YIELD.

Stage	Yield Genetic Architecture	H^2	Accuracy	Sensitivity	Specificity
PYT	3-categories QTL	0.20	0.90	0.52	0.95
		0.60	0.92	0.62	0.96
		1.00	0.94	0.72	0.97
	Uniform	0.20	0.92	0.59	0.95
		0.60	0.94	0.72	0.97
		1.00	0.94	0.69	0.97
	Negative Exponential	0.20	0.89	0.46	0.94
		0.60	0.90	0.52	0.95
		1.00	0.92	0.61	0.96
AYT-1	3-categories QTL	0.20	0.80	0.48	0.88
		0.60	0.88	0.70	0.93
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.80	0.49	0.88
		0.60	0.88	0.69	0.93
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.78	0.44	0.87
		0.60	0.87	0.67	0.92
		1.00	1.00	1.00	1.00
AYT-2	3-categories QTL	0.20	0.85	0.45	0.92
		0.60	0.91	0.65	0.95
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.87	0.50	0.92
		0.60	0.89	0.58	0.93
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.85	0.45	0.92
		0.60	0.91	0.65	0.95
		1.00	1.00	1.00	1.00
EYT	3-categories QTL	0.20	0.70	0.40	0.80
		0.60	0.90	0.80	0.93
		1.00	1.00	1.00	1.00
	Uniform	0.20	0.85	0.70	0.90
		0.60	0.85	0.70	0.90
		1.00	1.00	1.00	1.00
	Negative Exponential	0.20	0.80	0.60	0.87
		0.60	0.80	0.60	0.87
		1.00	1.00	1.00	1.00